

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocellaration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 5, 2006, 14:07:38 ; Search time 1757 Seconds  
(without alignments)  
582.346 Million cell updates/sec

Title: US-10-659-980a-5

Sequence: 18  
1 cccccctctcgcgatac 18

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 70 summaries

Database :

GenBml:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_htg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	18	6	CQ796912 Sequence
2	18	100.0	18	6	CS089158 Sequence
3	18	100.0	18	6	AX316096 Sequence
4	18	100.0	1457	3	AF386752 Nitroscemo
5	18	100.0	1457	3	AF386753 Nitroscemo
6	18	100.0	1457	6	CQ796908 Sequence
7	18	100.0	1457	6	CQ796909 Sequence
8	18	100.0	1457	6	CS089154 Sequence
9	18	100.0	1457	6	CS089155 Sequence
10	18	100.0	1457	6	AX316092 Sequence
11	18	100.0	1457	6	AX316093 Sequence
12	18	100.0	232644	14	AC098064 Rattus no
13	18	100.0	245255	14	AC121440 Rattus no
14	17	94.4	279	3	AJ853929 Unculture
15	17	94.4	471	3	AJ786628 Unculture
16	17	94.4	1426	3	AF386751 Nitroscemo
17	17	94.4	1445	1	AJ621032 Nitroscemo
18	17	94.4	1457	3	AF386757 Nitroscemo

#### ALIGNMENTS

C	19	17	94.4	1486	3	AB117705	AB117705 Unculture
C	20	16.4	91.1	18	6	CS089171	CS089171 Sequence
C	21	16.4	91.1	18	6	AX316109	AX316109 Sequence
C	22	16.4	91.1	1299	4	AB120381	AB120381 Equus cab
C	23	16.4	91.1	1378	4	AB120382	AB120382 Equus cab
C	24	16.4	91.1	1409	4	AB120380	AB120380 Equus cab
C	25	16.4	91.1	1422	1	AD16SRRN	X80412 A.denticole
C	26	16.4	91.1	73207	5	BX194202	BX194202 Zebrafish
C	27	16.4	91.1	81607	14	AC164449	AC164449 Bos tauru
C	28	16.4	91.1	100093	5	BX901896	BX901896 Zebrafish
C	29	16.4	91.1	115994	5	AL928940	AL928940 Mouse DNA
C	30	16.4	91.1	122848	5	CR457447	CR457447 Zebrafish
C	31	16.4	91.1	130639	8	AL356489	AL356489 Human DNA
C	32	16.4	91.1	141401	9	AL844585	AL844585 Mouse DNA
C	33	16.4	91.1	151256	14	CR940367	CR940367 Danio rer
C	34	16.4	91.1	156164	14	CR352215	CR352215 Danio rer
C	35	16.4	91.1	166031	14	CR933102	CR933102 Danio rer
C	36	16.4	91.1	166398	14	CR788228	CR788228 Danio rer
C	37	16.4	91.1	167473	14	AC145359	AC145359 Macropus
C	38	16.4	91.1	168373	9	BX530055	BX530055 Mouse DNA
C	39	16.4	91.1	174395	14	AC131875	AC131875 Rattus no
C	40	16.4	91.1	178531	14	CR792431	CR792431 Danio rer
C	41	16.4	91.1	180080	14	CR846078	CR846078 Danio rer
C	42	16.4	91.1	181839	14	CR589883	CR589883 Danio rer
C	43	16.4	91.1	182464	5	BX569779	BX569779 Zebrafish
C	44	16.4	91.1	185237	9	AL954390	AL954390 Mouse DNA
C	45	16.4	91.1	187216	5	CR376821	CR376821 Zebrafish
C	46	16.4	91.1	188181	9	AL591598	AL591598 Mouse DNA
C	47	16.4	91.1	191520	9	AL591598	AL591598 Mouse DNA
C	48	16.4	91.1	192040	14	AC153716	AC153716 Bos tauru
C	49	16.4	91.1	195247	5	BX322665	BX322665 Zebrafish
C	50	16.4	91.1	195220	14	AC105945	AC105945 Mus muscu
C	51	16.4	91.1	196033	5	CR759830	CR759830 Zebrafish
C	52	16.4	91.1	197999	14	AC117089	AC117089 Rattus no
C	53	16.4	91.1	197981	14	AC160068	AC160068 Bos tauru
C	54	16.4	91.1	205213	5	AL955173	AL955173 Zebrafish
C	55	16.4	91.1	213948	9	AC162927	AC162927 Mus muscu
C	56	16.4	91.1	216293	8	HTCRBCR9	HTCRBCR9 Homo sapi
C	57	16.4	91.1	216608	5	AL844559	AL844559 Zebrafish
C	58	16.4	91.1	222869	14	AC113777	AC113777 Rattus no
C	59	16.4	91.1	225149	14	AC103534	AC103534 Rattus no
C	60	16.4	91.1	226248	14	AC107501	AC107501 Rattus no
C	61	16.4	91.1	237632	14	AC150939	AC150939 Bos tauru
C	62	16.4	91.1	238937	14	AL773529	AL773529 Mus muscu
C	63	16.4	91.1	247670	14	AC117025	AC117025 Rattus no
C	64	16.4	91.1	251144	14	AC109565	AC109565 Rattus no
C	65	16.4	91.1	257087	14	AC152199	AC152199 Bos tauru
C	66	16.4	91.1	258549	14	AC133113	AC133113 Rattus no
C	67	16.4	91.1	258819	14	AC118292	AC118292 Rattus no
C	68	16.4	91.1	261202	14	AL929047	AL929047 Mus muscu
C	69	16.4	91.1	285603	14	AC104926	AC104926 Mus muscu
C	70	16.4	91.1	344922	14	CR376752	CR376752 Danio rer

RESULT 1  
CQ796912  
Sequence 5 from Patent WO2004026772.  
18 bp DNA  
Linear PAT 19-APR-2004  
LOCUS CQ796912  
DEFINITION Sequence 5 from Patent WO2004026772.  
ACCESSION CQ796912  
VERSION CQ796912.1 GI:46408538  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
other sequences; artificial sequences.  
REFERENCE  
1  
Hovanec, T.A.  
Ammunite-oxidizing bacteria and methods of using and detecting  
thesame  
Patent: WO 2004026772-A 5 01-APR-2004;  
JOURNAL Aquaria Inc. (US)

FEATURES  
source Location/Qualifiers

1..18  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide probe"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 18;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18  
|||||  
Db 1 CCCCCCTTCTGGATAC 18

RESULT 2  
LOCUS CS089158 18 bp DNA linear PAT 25-MAY-2005  
DEFINITION Sequence 5 from Patent EP1502948.  
ACCESSION CS089158  
VERSION CS089158.1 GI:67114442  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE  
AUTHORS Hovanec,T.A. and Burrell,P.C.  
TITLE Ammonia-oxidizing bacteria  
JOURNAL Patent: EP 1502948-A 5 02-FEB-2005;  
Aquadria Inc. (US)

FEATURES  
source Location/Qualifiers

1..18  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Description of Artificial Sequence: Probe"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 18;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18  
|||||  
Db 1 CCCCCCTTCTGGATAC 18

RESULT 3  
LOCUS AXJ16096 18 bp DNA linear PAT 14-DEC-2001  
DEFINITION Sequence 5 from Patent WO0190312.  
ACCESSION AXJ16096  
VERSION AXJ16096.1 GI:11789287  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE  
AUTHORS Hovanec,T.A. and Burrell,P.C.  
TITLE Ammonia-oxidizing bacteria  
JOURNAL Patent: WO 0190312-A 5 29-NOV-2001;  
AQUARIA, INC. (US)

FEATURES  
source Location/Qualifiers

1..18  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Probe"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 18;

Best Local Similarity 100.0%; Pred. No. 67;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18  
|||||  
Db 1 CCCCCCTTCTGGATAC 18

RESULT 4  
AF386752/c 1457 bp DNA linear ENV 10-MAY-2004  
LOCUS AF386752 Nitrosomonas sp. R7c131 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386752  
ACCESSION AF386752  
VERSION AF386752.1 GI:17864829  
KEYWORDS  
SOURCE Nitrosomonas sp. R7c131  
ORGANISM Nitrosomonas sp. R7c131

REFERENCE  
AUTHORS Burrell,P.C., Phalen,C.M. and Hovanec,T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in freshwater aquaria

JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 11722936  
2 (bases 1 to 1457)

REFERENCE  
AUTHORS Burrell,P.C., Phalen,C.M. and Hovanec,T.A.  
TITLE Direct Submission

JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria Group, 6100 Condon Dr, Moorpark, CA 93021, USA  
Location/Qualifiers

FEATURES  
source 1..1457  
/organism="Nitrosomonas sp. R7c131"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated from freshwater aquaria enrichments"  
/db\_xref="taxon:180511"  
/clone="R7c131"  
/environmental\_sample  
<1..>1457  
/product="16S ribosomal RNA"

rRNA

ORIGIN

Query Match 100.0%; Score 18; DB 3; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18  
|||||  
Db 118 CCCCCCTTCTGGATAC 101

RESULT 5  
AF386753 1457 bp DNA linear ENV 10-MAY-2004  
LOCUS AF386753 Nitrosomonas sp. R7c140 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386753  
ACCESSION AF386753  
VERSION AF386753.1 GI:17864830  
KEYWORDS  
SOURCE ENV.

ORGANISM Nitrosomonas sp. R7c140  
Nitrosomonas sp. R7c140  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;

REFERENCE  
AUTHORS Burrell,P.C., Phalen,C.M. and Hovanec,T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in freshwater aquaria

JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)

PUBMED 11722936

REFERENCE  
AUTHORS Burrell,P.C., Phalen,C.M. and Hovanec,T.A.  
TITLE Direct Submission

JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria



Group, 6100 Condor Dr, Moorpark, CA 93021, USA

FEATURES  
source  
1.1457  
/organism="Nitrosomonas sp. R7c140"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated from freshwater aquaria enrichments"  
/db\_xref="taxon:180543"  
/clone="R7c140"  
/environmental\_sample  
<1. >1457  
/product="16S ribosomal RNA"

ORIGIN  
RNA

Query Match  
Best Local Similarity 100.0%; Score 18; DB 3; Length 1457;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGGATAC 18  
118 CCCCCCTTCTGGATAC 101

Db

RESULT 6  
CO796908/c 1457 bp DNA linear PAT 19-APR-2004  
LOCUS  
DEFINITION Sequence 1 from Patent WO2004026772.  
ACCESSION CO796908  
VERSION CO796908.1 GI:46408534  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified  
FEATURES  
source  
1.1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="AOB Type A R7clone140 16S rDNA"

ORIGIN  
Location/Qualifiers

Query Match  
Best Local Similarity 100.0%; Score 18; DB 6; Length 1457;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGGATAC 18  
118 CCCCCCTTCTGGATAC 101

Db

RESULT 7  
CO796909/c 1457 bp DNA linear PAT 19-APR-2004  
LOCUS  
DEFINITION Sequence 2 from Patent WO2004026772.  
ACCESSION CO796909  
VERSION CO796909.1 GI:46408535  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified  
FEATURES  
source  
1.1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="Description of Unknown Organism: AOB Type A1  
ammonia-oxidizing bacterium represented by R7 clone187"

ORIGIN  
Location/Qualifiers

source  
1.1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="AOB Type A1 R7clone187 16S rDNA"

ORIGIN  
Location/Qualifiers

Query Match  
Best Local Similarity 100.0%; Score 18; DB 6; Length 1457;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGGATAC 18  
118 CCCCCCTTCTGGATAC 101

Db

RESULT 8  
CS089154/c 1457 bp DNA linear PAT 25-MAY-2005  
LOCUS  
DEFINITION Sequence 1 from Patent EP1502948.  
ACCESSION CS089154  
VERSION CS089154.1 GI:66714438  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified  
FEATURES  
source  
1.1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="Description of Unknown Organism: AOB Type A  
ammonia-oxidizing bacterium represented by R7 clone140"

ORIGIN  
Location/Qualifiers

Query Match  
Best Local Similarity 100.0%; Score 18; DB 6; Length 1457;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGGATAC 18  
118 CCCCCCTTCTGGATAC 101

Db

RESULT 9  
CS089155/c 1457 bp DNA linear PAT 25-MAY-2005  
LOCUS  
DEFINITION Sequence 2 from Patent EP1502948.  
ACCESSION CS089155  
VERSION CS089155.1 GI:66714439  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified  
FEATURES  
source  
1.1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="Description of Unknown Organism: AOB Type A1  
ammonia-oxidizing bacterium represented by R7 clone187"

ORIGIN  
Location/Qualifiers

Query Match 100.0%; Score 18; DB 6; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCCCTCTTGATAC 18  
Db 118 CCCCCCTCTTGATAC 101

RESULT 10  
AX316092/c 1457 bp DNA linear PAT 14-DEC-2001  
LOCUS Sequence 1 from Patent WO0190312.  
DEFINITION AX316092  
ACCESSION AX316092.1 GI:17899283  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1.1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="AOB Type A ammonia-oxidizing bacterium represented by R7 clone140"

ORIGIN  
Query Match 100.0%; Score 18; DB 6; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCCCTCTTGATAC 18  
Db 118 CCCCCCTCTTGATAC 101

RESULT 11  
AX316093/c 1457 bp DNA linear PAT 14-DEC-2001  
LOCUS Sequence 2 from Patent WO0190312.  
DEFINITION AX316093  
ACCESSION AX316093.1 GI:17899284  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1.1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="AOB Type A1 ammonia-oxidizing bacterium represented by R7 clone187"

ORIGIN  
Query Match 100.0%; Score 18; DB 6; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCCCTCTTGATAC 18  
Db 118 CCCCCCTCTTGATAC 101

RESULT 12  
AC098064  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE

AC098064 232644 bp DNA linear HTG 14-SEP-2004  
Rattus norvegicus clone CH230-117A23, WORKING DRAFT SEQUENCE, 2  
ordered pieces.  
AC098064  
AC098064.6 GI:5200538  
HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Murinae; Rattus.  
1 (bases 1 to 232644)  
Muzny, D., Marle, Metzger, M., Lee, Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D.,  
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, Y., Carter, K., Cavazos, I., Cesaar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Dayla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebreyes, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,  
Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpachy, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorensuhewa, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,  
Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,  
Morris, S., Muidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,  
Newton, N., Nguyen, N., Norris, S., Nwakoelam, O., Okunolu, G.,  
Olanunwasagun, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,  
Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A.,  
Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E.,  
Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,  
Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojda, A.,  
Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Shetter, S.,  
Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A.,  
Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E.,  
Song, X.-Z., Sorrelle, R., Soza, J., Steimle, M., Strong, R., Sutton, A.,  
Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,  
Tingey, A., Trejo, Z., Usmani, K., Vals, J., Wang, S., Warren, S.,  
Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, D.,  
Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleczek, R.,  
Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,  
Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,  
Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R.,  
Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE

2 (bases 1 to 232644)  
Worley, K. C.  
Direct Submission  
Submitted (23-Oct-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 232644)



## COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Sep 19, 2002 this sequence version replaced gi:21908223.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

## ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

## ----- Project Information

Center project name: GX00

Center clone name: CH230-16F9

## ----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 229731 bases at least Q40

Consensus quality: 232074 bases at least Q30

Consensus quality: 233846 bases at least Q20

Estimated insert size: 258987; sum-of-contigs estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

## -----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html))  
\* NOTE: This sequence may represent more than one clone.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 99258: contig of 99258 bp in length

\* 99259 99358: gap of unknown length

\* 99359 245255: contig of 145897 bp in length.

## FEATURES

source

1. 245255

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-16F9"

1254. .2026

/note="clone\_boundary"

clone\_end:T7

misc\_feature

site:BCORI

end\_sequence:BH341961"

1967. .2402

/note="clone boundary"

clone\_end:Sp6

misc\_feature

site:BCORI

end\_sequence:BH341989"

99259. .99358

/estimated\_length=unknown

155481. .156840

/note="wgs\_end\_extension"

clone\_end:Sp6"

## ORIGIN

Query Match

Best Local Similarity 100.0%; Score 18; DB 14; Length 245255;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18

DB 141992 CCCCCCTCTTCGTGATAC 142009

RESULT 14

AJ853929/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

gene  
rRNA  
/country="Belgium"  
1..471  
/gene="16S rRNA"  
<1..471  
/gene="16S rRNA"  
/product="16S ribosomal RNA"

## ORIGIN

Query Match 94.4%; Score 17; DB 3; Length 471;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCCTCTTCGATAC 18  
|||||  
Db 137 CCCCTCTTCGATAC 121

RESULT 16  
AF386751/c 1426 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosomonas sp. R7c155 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386751  
ACCESSION AF386751 GI:17864828  
VERSION AF386751.1 GI:17864828  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. R7c155  
ORGANISM Nitrosomonas sp. R7c155  
REFERENCE Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas; environmental samples.  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 11722936  
REFERENCE 2 (bases 1 to 1426)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condon Dr, Moorpark, CA 93021, USA  
LOCATION/Qualifiers

## FEATURES

Source  
1..1426  
/organism="Nitrosomonas sp. R7c155"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichments"  
/db\_xref="taxon:180510"  
/clone="R7c155"  
/environmental\_sample  
<1..>1426  
/product="16S ribosomal RNA"

## ORIGIN

Query Match 94.4%; Score 17; DB 3; Length 1426;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCCTCTTCGATAC 18  
|||||  
Db 86 CCCCTCTTCGATAC 70

RESULT 17  
AJ621032/c 1445 bp DNA linear BCT 08-JUN-2004  
LOCUS Nitrosomonas sp. Is343 16S rRNA gene, isolate Is343.  
DEFINITION AJ621032  
ACCESSION AJ621032.1 GI:4094851  
VERSION AJ621032.1 GI:4094851  
KEYWORDS 16S ribosomal RNA; 16S rRNA gene.  
SOURCE Nitrosomonas sp. Is343  
ORGANISM Nitrosomonas sp. Is343  
REFERENCE Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1

AUTHORS Bollmann, A., Schmidt, I. and Bodelier, P.L.E.  
TITLE Influence of salt on activity, growth, protein pattern and fatty  
acid composition of the new isolated ammonia oxidizing strain  
Nitrosomonas strain Is343

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1445)  
AUTHORS Bollmann, A.

JOURNAL Direct Submission  
Submitted (15-JAN-2004) Bollmann A., Department for Microbial  
Ecology, NIOO-KNAW Centre for Limnology, Rijksstraatweg 6, 3631 AC  
Neuversluis, NETHERLANDS  
LOCATION/Qualifiers

## FEATURES

Source  
1..1445  
/organism="Nitrosomonas sp. Is343"  
/mol\_type="genomic DNA"  
/isolate="Is343"  
/isolation\_source="brakish water"  
/db\_xref="taxon:261298"  
/country="Netherlands:River Scheide"  
1..144  
/gene="16S rRNA"  
1..144  
/gene="16S rRNA"  
/product="16S ribosomal RNA"

## ORIGIN

Query Match 94.4%; Score 17; DB 1; Length 1445;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCCTCTTCGATAC 18  
|||||  
Db 113 CCCCTCTTCGATAC 97

RESULT 18  
AF386757/c 1457 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosomonas sp. R7c187 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386757  
ACCESSION AF386757 GI:17864834  
VERSION AF386757.1 GI:17864834  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. R7c187  
ORGANISM Nitrosomonas sp. R7c187  
REFERENCE Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas; environmental samples.  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 11722936  
REFERENCE 2 (bases 1 to 1457)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condon Dr, Moorpark, CA 93021, USA  
LOCATION/Qualifiers

## FEATURES

Source  
1..1457  
/organism="Nitrosomonas sp. R7c187"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichments"  
/db\_xref="taxon:180512"  
/clone="R7c187"  
/environmental\_sample  
<1..>1457  
/product="16S ribosomal RNA"

## ORIGIN

Query Match 94.4%; Score 17; DB 3; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCCCCTCTTGATAC 18  
|||||  
DB 117 CCCCCCTCTTGATAC 101

RESULT 19  
LOCUS AB117705/c 1486 bp DNA linear ENV 27-JUL-2005  
DEFINITION Uncultured bacterium gene for 16S ribosomal RNA, partial sequence,  
clone:NB-01.  
ACCESSION AB117705  
VERSION AB117705.1 GI:42627690  
KEYWORDS ENV.  
SOURCE uncultured bacterium  
ORGANISM uncultured bacterium  
REFERENCE 1  
AUTHORS Kindaichi,T., Ito,T. and Okabe,S.  
TITLE Ecophysiological interaction between nitrifying bacteria and heterotrophic bacteria in autotrophic nitrifying biofilms as determined by microautoradiography-fluorescence in situ hybridization  
JOURNAL Appl. Environ. Microbiol. 70 (3), 1641-1650 (2004)  
PUBMED 1506789  
REFERENCE 2 (bases 1 to 1486)  
AUTHORS Kindaichi,T.  
TITLE Direct Submision  
JOURNAL Submitted (18-AUG-2003) Tomonori Kindaichi, Hokkaido university, Graduate school of engineering, Department of urban and environmental engineering, Kita 13, Nishi 8, Kita-Ku, Sapporo, Hokkaido, 060-8628, Japan (E-mail:kindeng.hokudai.ac.jp, Tel:81-11-706-6267, Fax:81-11-706-6267)  
Location/Qualifiers  
FEATURES  
source  
1. .1486  
/organism="uncultured bacterium"  
/mol\_type="genomic DNA"  
/isolation\_source="autotrophic nitrifying biofilm"  
/db\_xref="taxon:77133"  
/clone="NB-01"  
/environmental\_sample  
<1. .>1486  
/product="16S ribosomal RNA"

ORIGIN  
RNA  
Query Match 94.4%; Score 17; DB 3; Length 1486;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCCCCTCTTGATAC 18  
|||||  
DB 134 CCCCCCTCTTGATAC 118

RESULT 20  
LOCUS CS089171 18 bp DNA linear PAT 25-MAY-2005  
DEFINITION Sequence 18 from Patent EP1502948.  
ACCESSION CS089171  
VERSION CS089171.1 GI:6714455  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Hovanec,T.A. and Burrell,P.C.  
TITLE Ammonia-oxidizing bacteria  
JOURNAL Patent: EP 1502948-A 18 02-FEB-2005;  
Aguarria Inc. (US)  
Location/Qualifiers  
FEATURES  
source  
1. .18  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"

ORIGIN  
/db\_xref="taxon:32630"  
/note="Description of Artificial Sequence: Primer"

Query Match 91.1%; Score 16.4; DB 6; Length 18;  
Best Local Similarity 94.4%; Pred. No. 5.3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATAC 18  
|||||  
DB 1 CCCCCCTCTTGACTAC 18

RESULT 21  
LOCUS AX316109 18 bp DNA linear PAT 14-DEC-2001  
DEFINITION Sequence 18 from Patent WO0190312.  
ACCESSION AX316109  
VERSION AX316109.1 GI:17899300  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Hovanec,T.A. and Burrell,P.C.  
TITLE Ammonia-oxidizing bacteria  
JOURNAL Patent: WO 0190312-A 18 29-NOV-2001;  
AQUARIA, INC. (US)  
Location/Qualifiers  
FEATURES  
source  
1. .18  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Primer"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 6; Length 18;  
Best Local Similarity 94.4%; Pred. No. 5.3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATAC 18  
|||||  
DB 1 CCCCCCTCTTGACTAC 18

RESULT 22  
LOCUS AB120381 1299 bp mRNA linear MAM 26-JUL-2005  
DEFINITION Equus caballus Ly49B mRNA for killer cell lectin-like receptor, complete cds, allele:Ly49B\*002.  
ACCESSION AB120381  
VERSION AB120381.1 GI:45597240  
KEYWORDS  
SOURCE Equus caballus (horse)  
ORGANISM Equus caballus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.  
REFERENCE 1  
AUTHORS Takahashi,T., Yawata,M., Raudepp,T., Lear,T.L., Chowdhary,B.P., Antczak,D.F. and Kasahara,M.  
TITLE Natural killer cell receptors in the horse: evidence for the existence of multiple transcribed Ly49 genes  
JOURNAL Eur. J. Immunol. 34 (3), 773-784 (2004)  
PUBMED 14991607  
REFERENCE 2 (bases 1 to 1299)  
AUTHORS Takahashi,T. and Kasahara,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-SEP-2003) Masanori Kasahara, Hokkaido University Graduate School of Medicine, Department of Pathology, North-15, West-7, Sapporo, 060-0638, Japan (E-mail:m.kasahara@med.hokudai.ac.jp, Tel:81-11-706-5048, Fax:81-11-706-7825)  
Location/Qualifiers  
FEATURES  
source  
1. .1299

```

/gene="Ly49B"
1. .1295
/gene="Ly49B"
1. .50
/gene="Ly49B"
51. .899
/gene="Ly49B"
/codon_start=1
/evidence=experimental
/allele="Ly49B*002"
/product="killer cell lectin-like receptor"
/protein_id="BAD1281.1"
/db_xref="GI:45597241"
/training="MSSEVYCSLRFPSPLSEQNTLRPGTORPEKTDKEFVPM
HPFVATGILLCLLVTYVIGMIPOCFEKKHOEEILRLIPQYSINNNYSEKQ
LILNKTILACDILKRLTHQKEKDLSPNEKKRCHIKOELPSKLSQSGKNEHMSCC
GVNCFYFTSENENMGCEKCFORYNLSLKTDDEKLFVQRTIRSYIGLSYNT
ESKKMWIDSGMSGINFKIVSLPSGRGRGKAFISSIRIADICAKKYKICEKTD
FTACFN"
900. .1299
/gene="Ly49B"

3' UTR

ORIGIN
Query Match 91.1%; Score 16.4; DB 4; Length 1299;
Best Local Similarity 94.4%; Pred. No. 5.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATAC 18
|||||
|
Db 1223 CACCCTCTCTGATAC 1240

RESULT 23
ABI20382 1378 bp mRNA linear MAM 26-JUL-2005
LOCUS Equus caballus Ly49B mRNA for killer cell lectin-like receptor,
DEFINITION complete cds, allele:Ly49B-0010102 variant 1.
ACCESSION ABI20382
VERSION ABI20382.1 GI:45597242
KEYWORDS
SOURCE
ORGANISM Equus caballus (horse)
EQUUS caballus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.
REFERENCE
AUTHORS 1 Takahashi,T., Yawata,M., Raudepp,T., Lear,T.L., Chowdhary,B.P.,
Antczak,D.F. and Kasahara,M.
TITLE Natural killer cell receptors in the horse: evidence for the
existence of multiple transcribed Ly49 genes
JOURNAL Eur. J. Immunol. 34 (3), 773-784 (2004)
PUBMED 14991607
REFERENCE 2 (bases 1 to 1378)
AUTHORS Takahashi,T. and Kasahara,M.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-2003) Masanori Kasahara, Hokkaido University,
Graduate School of Medicine, Department of Pathology, North-15,
West-7, Sapporo, 060-0638, Japan (E-mail:mkasahamed.hokudai.ac.jp)
Tel:81-11-706-5048, Fax:81-11-706-7825)
FEATURES
source
1. .1378
/organism="Equus caballus"
/mol_type="mRNA"
/isolate="Horse#2"
/db_xref="taxon:9796"
/feature_type="spleen"
1. .1378
/gene="Ly49B"
1. .112
/gene="Ly49B"
5' UTR

```

```

113..853
/gene="Ly49B"
/codon_start=1
/evidence=experimental
/allele="Ly49B*0010102 variant 1"
/product="killer cell lectin-like receptor"
/protein_id="BAD12812.1"
/db_xref="GI:45597243"
/translation="MSSQEVYCSLARPLQSLRQNTLRPGNRPCKTDKESVPM
HPVATVGLICLLLVTVLVGTMITQCFEKKQOEILRLKLPQYDSIRNNYSKEQ
LNLINKLECDILNKNKTLHQKELDSLPEMKRCHIKOELPSKSLQSGKFNEDMSCC
GVNVCITSENNEMKGTCKEKCESYNLSLKLIDEDKLNFPVQRTYRDSYWGISYNAT
ESKMKMIDSGMSSGIFTHFQADVILIK"
B54..1378
/gene="Ly49B"

3' UTR

ORIGIN
Query Match 91.1%; Score 16.4; DB 4; Length 1378;
Best Local Similarity 94.4%; Pred. No. 5.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGGATAC 18
|||||
Db 1314 CACCCCTCTCTGGATAC 1331

RESULT 24
ABI20380 1409 bp mRNA linear MAM 26-JUL-2005
ABI20380 Equus caballus Ly49B mRNA for killer cell lectin-like receptor,
LOCUS complete cds, allele:Ly49B*0010101.
ACCESSION ABI20380
VERSION ABI20380.1 GI:45597238
KEYWORDS
SOURCE
ORGANISM
EQUUS caballus (horse)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.
REFERENCE
AUTHORS 1 Takahashi,T., Ywata,M., Raudsepp,T., Lear,T.L., Chowdhary,B.P.,
Antczak,D.F. and Kasahara,M.
TITLE Natural killer cell receptors in the horse: evidence for the
existence of multiple transcribed Ly49 genes
Eur. J. Immunol. 34 (3), 773-784 (2004)
14591607
2 (bases 1 to 1409)
Takahashi,T. and Kasahara,M.
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (16-SEP-2003) Masanori Kasahara, Hokkaido University
Graduate School of Medicine, Department of Pathology, North-15,
West-7, Sapporo, 060-0638, Japan (E-mail:mkasahamed.hokudai.ac.jp,
Tel:81-11-706-5048, Fax:81-11-706-7825)
FEATURES
Source Location/Qualifiers
1..1409
/organism="Equus caballus"
/mol_type="mRNA"
/isolate="Horse#1"
/db_xref="taxon:9796"
/tissue_type="spleen"
1..1409
/gene="Ly49B"
1..149
/gene="Ly49B"
150..998
/gene="Ly49B"
/codon_start=1
/evidence=experimental
/allele="Ly49B*0010101"
/product="killer cell lectin-like receptor"
/protein_id="BAD12810.1"
/db_xref="GI:45597239"
/translation="MSSQEVYCSLARPLQSLRQNTLRPGNRPCKTDKESVPM
HPVATVGLICLLLVTVLVGTMITQCFEKKQOEILRLKLPQYDSIRNNYSKEQ

```







gap /estimated\_length=unknown  
30566.30665  
gap /estimated\_length=unknown  
37258.37357  
gap /estimated\_length=unknown  
39528.39627  
gap /estimated\_length=unknown  
42301.42400  
gap /estimated\_length=unknown  
49882.49981  
gap /estimated\_length=unknown  
52350.52449  
gap /estimated\_length=unknown  
54994.55093  
gap /estimated\_length=unknown  
71342.71441  
gap /estimated\_length=unknown  
73527.73626  
gap /estimated\_length=unknown  
75724.75823  
gap /estimated\_length=unknown  
78618.78717  
gap /estimated\_length=unknown

## ORIGIN

Query Match 91.1%; Score 16.4; DB 14; Length 81607;  
Best Local Similarity 94.4%; Pred. No. 5.7e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGATAC 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1127 CACCCCTCTTCGATAC 1144

RESULT 28  
BX901896 100093 bp DNA linear VRT 22-JAN-2005  
LOCUS zebrafish DNA sequence from clone DKEY-71L1 in linkage group 5,  
DEFINITION complete sequence.  
ACCESSION BX901896 GI:58036866  
VERSION BX901896.7 GI:58036866  
KEYWORDS HTG.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio

## SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 100093)  
Holt, K.

REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (22-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,  
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Jan 22, 2005 this sequence version replaced gi:57337757.

## COMMENT

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [zf1sh-help@sanger.ac.uk](mailto:zf1sh-help@sanger.ac.uk)  
-----  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare

occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information  
on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) Clone-derived  
zebrafish pUC subclones occasionally display inconsistency over the  
length of mononucleotide A/T runs and conserved TA repeats. Where  
this is found the longest good quality representation will be  
submitted.  
Repeat names beginning 'Dr' were identified by the Recon repeat  
discovery system (Zhifeng Bao and Sean Eddy, submitted), and those  
beginning 'drr' were identified by Rick Waterman (Stephen Johnson  
lab, WashU). For further information see  
[http://www.sanger.ac.uk/Projects/D\\_rerio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml) DKEY-71L1  
is from a Zebrafish BAC library  
VECTOR: pindigobAC-5.  
Location/Qualifiers  
1..100093  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEY-71L1"  
/clone\_1ib="DanioKey"

## FEATURES

source

## ORIGIN

Query Match 91.1%; Score 16.4; DB 5; Length 100093;  
Best Local Similarity 94.4%; Pred. No. 5.7e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGATAC 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 18254 CCCCCCTCTTCGATAC 18271

RESULT 29  
AL928940 115994 bp DNA linear ROD 15-JUL-2003  
LOCUS Mouse DNA sequence from clone RP23-9A017 on chromosome 2, complete  
DEFINITION sequence.  
ACCESSION AL928940 GI:32812763  
VERSION AL928940.12 GI:32812763  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

## SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 115994)  
Sycamore, N.

REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (15-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,  
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Jul 15, 2003 this sequence version replaced gi:32567729.

## COMMENT

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
-----  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-94J17 is from the RPCI-23 Mouse BAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.6.

#### FEATURES

source

1..115994  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="2"  
/clone="RP23-94J17"  
/clone\_1ib="RPCI-23"

#### ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 115994;  
Best Local Similarity 94.4%; Pred. No. 5.7e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCCCCCTCTTGGATAC 18  
Db 4106 CCCCCCTCTTGGATAC 4123

RESULT 30  
LOCUS CR457447 122848 bp DNA linear VRT 02-MAR-2005  
DEFINITION zebrafish DNA sequence from clone CH211-108P6 in linkage group 15, complete sequence.  
ACCESSION CR457447  
VERSION CR457447.5 GI:60458002  
KEYWORDS HTG.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 122848)  
Phillimore, B.  
Direct Submission  
Submitted (02-MAR-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk  
On Mar 2, 2005 this sequence version replaced gi:58578408.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare

occasion of the clone being a YAC. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhong Bao and Sean Bddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see [http://www.sanger.ac.uk/Projects/D\\_rerio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml) CH211-108P6 is from a CHORI-211 BAC library VECTOR: pFARBAC2.1.

#### FEATURES

source

1..122848  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="CH211-108P6"  
/clone\_1ib="CHORI-211"

#### ORIGIN

Query Match 91.1%; Score 16.4; DB 5; Length 122848;  
Best Local Similarity 94.4%; Pred. No. 5.7e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCCCCCTCTTGGATAC 18  
Db 44073 CCCCCCTCTTGGATAC 44090

RESULT 31  
LOCUS AL356489/c 130639 bp DNA linear PRI 27-JUL-2005  
DEFINITION Human DNA sequence from clone RP11-384P7 on chromosome 9 containing the TRBV25OR9-2 gene for T cell receptor beta variable 25/OR9-2, the PTENP1 gene for phosphatase and tensin homolog (mutated in multiple advanced cancers 1) pseudogene 1 (PTN2, PTEN2, PSIPEN1), a T-cell receptor beta chain V region protein pseudogene, two novel genes, the 5' end of the PRSS3 gene for serine protease 3 (mesotrypsin) and two CpG islands, complete sequence.  
ACCESSION AL356489  
VERSION AL356489.14 GI:14547259  
KEYWORDS HTG; CpG island; mesotrypsin; MTG; protease; PRSS3; PSIPEN1; PTEN2; PTENP1; T cell receptor; TCRBV110; TRBV25/OR9-2; TRY4; tryptsin.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 130639)  
Laird, G.  
Direct Submission  
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: veg@sanger.ac.uk  
On Jun 25, 2001 this sequence version replaced gi:14270561. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9> R11-384P7 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see

#### COMMENT



```

misc_feature
/locus_tag="RP11-176F3.3-002"
join(95630..95667,AL358573.25:11047..11174,
AL358573.25:12944..13103,AL358573.25:14130..14383,
AL358573.25:14787..>14829)
/feature_key="PRSS3"
/locus_tag="RP11-176F3.3-002"
/product="protease, serine, 3 (mesotrypsin)"
/feature_key="match: ESTs: BF312188"
130639
/feature_key="Clone_right_end: RP11-384P7"

ORIGIN
Query Match          91.1%; Score 16.4; DB 8; Length 130639;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18
    |||||
Db 85401 CTCCTCTTCGTGATAC 85384

RESULT 32
AL844585 141401 bp DNA linear ROD 28-FEB-2003
LOCUS Mouse DNA sequence from clone RP23-33915 on chromosome 4, complete
DEFINITION
ACCESSION AL844585
VERSION AL844585.11 GI:28625371
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 141401)
Andrew.R.
Direct Submission
Submitted (28-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk
On Feb 28, 2003 this sequence version replaced gi:286253720.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Emi., EMBL; Sw., SWISSPROT; Tr., TREMBL; Mp., MORMEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.choi.org/bacpac/home.htm

```

```

FEATURES
source
/locus_tag="PBAC3.6"
Location/Qualifiers
1..141401
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-33915"
/clone_11b="RPCT-23"

ORIGIN
Query Match          91.1%; Score 16.4; DB 9; Length 141401;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18
    |||||
Db 27371 CTCCTCTTCGTGATAC 27354

RESULT 33
CR940367 151256 bp DNA linear HTG 14-JUL-2005
LOCUS Danio rerio chromosome 12 clone DKEX-245K17, WORKING DRAFT
DEFINITION
ACCESSION CR940367
VERSION CR940367.5 GI:70908015
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 151256)
McLaren, S.
Direct Submission
Submitted (12-JUL-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk
http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#date1htc
On Jul 15, 2005 this sequence version replaced gi:62460938.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
Project Information
Center project name: ZK245K17
Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 149014 bases at least Q40
Consensus quality: 149382 bases at least Q30
Consensus quality: 149654 bases at least Q20
Insert size: 150856; sum-of-coverage
Insert size: 154858; 0.8% error; agrose-fp
Quality coverage: 7.88x in Q20 bases; sum-of-coverage
Quality coverage: 7.68x in Q20 bases; agrose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 10458: contig of 10458 bp in length
* 10459: gap of 100 bp
* 10559: contig of 20720 bp in length
* 31279: gap of 100 bp
* 31379: contig of 40580 bp in length
* 71958: gap of 100 bp
* 72058: gap of 100 bp

```

```

* 72059 144791: contig of 72733 bp in length
* 144792 144891: gap of 100 bp
* 144892 151256: contig of 6365 bp in length.
FEATURES
    source
        1. 151256
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /chromosome="12"
            /clone="DKEY-245K17"
            /clone_1ib="DantolKey"
            1. 10458
            /note="assembly_fragment:01537
            fragment_chain:1"
            10559. 31278
            /note="assembly_fragment:01640
            fragment_chain:1"
            31379. 71958
            /note="assembly_fragment:00001"
            72059. 144791
            /note="assembly_fragment:00531"
            144892. 151256
            /note="assembly_fragment:01460.0"
ORIGIN
Query Match      91.1%; Score 16.4; DB 14; Length 151256;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CCCCCCTTCTTGATAC 18
Db      133432 CCCCCTTCTTGATAC 133415

```

```

RESULT 34
CR352215      156164 bp      DNA      linear      HTG 10-AUG-2005
LOCUS      Danio rerio chromosome 1 clone CH211-235F1, WORKING DRAFT SEQUENCE.
ACCESSION      CR352215.24 GI:72420623
VERSION      HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS      Danio rerio (zebrafish)
SOURCE      Danio rerio
ORGANISM      Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
            1 (bases 1 to 156164)
REFERENCE      Oliver K.
AUTHORS      Direct Submission
TITLE      Submitted (09-AUG-2005) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL      Cambridgehire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests:
            http://www.sanger.ac.uk/Projects/D_rerio/fage.shtml#dataeight
COMMENT      On Aug 12, 2005 this sequence version replaced gi:70907895.
            ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            ----- Project Information
            Center project name: zC235F1
            ----- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 155692 bases at least Q40
            Consensus quality: 155821 bases at least Q30
            Consensus quality: 155898 bases at least Q20
            Insert size: 156164; sum-of-contigs
            Insert size: 146349; 16.2% error; agarose-fp
            Quality coverage: 13.63x in Q20 bases; sum-of-contigs Quality
            coverage: 14.79x in Q20 bases; agarose-fp
            -----
            * NOTE: This is a 'working draft' sequence. It currently

```

```

* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 156164: contig of 156164 bp in length.
FEATURES
    source
        1. 156164
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /chromosome="1"
            /clone="CH211-235F1"
            /clone_1ib="CHORI-211"
            1. 156164
            /note="assembly_fragment:01391
            clone_end:r7
            vector_side:right"
ORIGIN
Query Match      91.1%; Score 16.4; DB 14; Length 156164;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CCCCCCTTCTTGATAC 18
Db      128074 CCCCCTTCTTGATAC 128091

```

```

RESULT 35
CR933102      166031 bp      DNA      linear      HTG 22-MAY-2005
LOCUS      Danio rerio clone DKEY-256H11, WORKING DRAFT SEQUENCE, 5 unordered
DEFINITION      pieces.
ACCESSION      CR933102
VERSION      CR933102.3 GI:66392828
KEYWORDS      HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
            1 (bases 1 to 166031)
REFERENCE      McLaren,S.
AUTHORS      Direct Submission
TITLE      Submitted (21-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL      Cambridgehire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT      On May 22, 2005 this sequence version replaced gi:57863685.
            ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            ----- Project Information
            Center project name: zK256H11
            ----- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 163592 bases at least Q40
            Consensus quality: 164014 bases at least Q30
            Consensus quality: 164549 bases at least Q20
            Insert size: 165631; sum-of-contigs
            Insert size: 165725; 1.7% error; agarose-fp
            Quality coverage: 7.32x in Q20 bases; sum-of-contigs Quality
            coverage: 7.93x in Q20 bases; agarose-fp
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is

```

\* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 2448: contig of 2448 bp in length  
 \* 2449 2548: gap of 100 bp  
 \* 2549 6036: contig of 3488 bp in length  
 \* 6037 6136: gap of 100 bp  
 \* 6137 42611: contig of 36475 bp in length  
 \* 42612 42712: gap of 100 bp  
 \* 42712 98617: contig of 55906 bp in length  
 \* 98618 98717: gap of 100 bp  
 \* 98718 166031: contig of 67314 bp in length.  
 Location/Qualifiers

FEATURES  
 source

1..166031  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone\_id="DKEY-256H11"  
 /clone\_lib="DantioKey"  
 1..2448  
 /note="assembly\_fragment:00034  
 fragment\_chain:1"  
 2549..6036  
 /note="assembly\_fragment:00047  
 fragment\_chain:1"  
 6137..42611  
 /note="assembly\_fragment:00070  
 fragment\_chain:2"  
 42712..98617  
 /note="assembly\_fragment:00482  
 fragment\_chain:2"  
 98718..166031  
 /note="assembly\_fragment:01206.0"

ORIGIN  
 Query Match 91.1%; Score 16.4; DB 14; Length 166031;  
 Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGATAC 18  
 |||||  
 Db 60600 CCCCCCTTCTGATAC 60617

RESULT 36  
 CR788228/c CR788228 166398 bp DNA linear HTG 05-AUG-2005  
 LOCUS Danio rerio chromosome 3 clone CH211-15N4, WORKING DRAFT SEQUENCE,  
 DEFINITION 9 unordered pieces.  
 ACCSSION CR788228.6 GI:71891505  
 VERSION HTG; HTGS PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
 KEYWORDS Danio rerio (zebrafish)  
 SOURCE Danio rerio  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Osteichthyes;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 166398)  
 Dunn, M.

REFERENCE  
 AUTHORS Direct Submission  
 TITLE Submitted (04-AUG-2005) Wellcome Trust Sanger Institute, Hinxton,  
 JOURNAL Cambridgehire, CB10 1SA, UK. E-mail enquiries:  
 zfish-help@sanger.ac.uk Clone requests:  
 http://www.sanger.ac.uk/projects/D-rerio/fsgs.shtml#dataight  
 On Aug 5, 2005 this sequence version replaced gi:62867845.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: zfish-help@sanger.ac.uk  
 ----- Project Information

Center project name: zc15N4  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Chemistry: Dye-terminator; 100% of reads  
 Consensus quality: 163105 bases at least Q40  
 Consensus quality: 163661 bases at least Q30  
 Consensus quality: 164201 bases at least Q20  
 Insert size: 165598; sum-of-contigs  
 Insert size: 144488; 6.1% error; agarose-fp  
 Quality coverage: 9.85x in Q20 bases; sum-of-contigs Quality  
 coverage: 11.29x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 9 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 58743: contig of 58743 bp in length  
 \* 58744 58843: gap of 100 bp  
 \* 58844 67253: contig of 8410 bp in length  
 \* 67254 67353: gap of 100 bp  
 \* 67354 73023: contig of 5670 bp in length  
 \* 73024 73123: gap of 100 bp  
 \* 73124 78003: contig of 4880 bp in length  
 \* 78004 78103: gap of 100 bp  
 \* 78104 86607: contig of 8504 bp in length  
 \* 86608 86707: gap of 100 bp  
 \* 86708 89095: contig of 2388 bp in length  
 \* 89096 89195: gap of 100 bp  
 \* 89196 92090: contig of 2895 bp in length  
 \* 92091 92190: gap of 100 bp  
 \* 92191 107680: contig of 15490 bp in length  
 \* 107681 107781: gap of 100 bp  
 \* 107781 166398: contig of 58618 bp in length.  
 Location/Qualifiers

FEATURES  
 source

1..166398  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /chromosome="3"  
 /clone="CH211-15N4"  
 /clone\_lib="CHORI-211"  
 1..58743  
 /note="assembly\_fragment:01578  
 fragment\_chain:1"  
 58844..67253  
 /note="assembly\_fragment:00232  
 fragment\_chain:1"  
 67354..73023  
 /note="assembly\_fragment:00179  
 fragment\_chain:1"  
 73124..78003  
 /note="assembly\_fragment:00623  
 fragment\_chain:2"  
 78104..86607  
 /note="assembly\_fragment:00391  
 fragment\_chain:2"  
 86708..89095  
 /note="assembly\_fragment:00002"  
 89196..92090  
 /note="assembly\_fragment:00365  
 fragment\_chain:3"  
 92191..107680  
 /note="assembly\_fragment:00280  
 fragment\_chain:3"  
 107781..166398  
 /note="assembly\_fragment:01579  
 fragment\_chain:3  
 clone end: T7  
 vector\_side:right"

ORIGIN

Query Match 91.1%; Score 16.4; DB 14; Length 166398;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18  
|||||

Db 32744 CCCCCCTCTTCGTGATAC 32727

RESULT 37  
AC145359/c 167423 bp DNA linear HTG 19-JUL-2003  
LOCUS Macropus eugenii clone ME\_KBa-138F17, WORKING DRAFT SEQUENCE, 5  
DEFINITION  
AC145359  
AC145359  
AC145359.2 GI:32996760  
HTG; HTGS\_PHASE2; HTGS\_DRAFT  
KEYWORDS  
SOURCE Macropus eugenii (tammar wallaby)  
ORGANISM Macropus eugenii  
Macropus eugenii  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Metatheria; Diprodontia; Macropodidae; Macropus.  
1 (bases 1 to 167423)  
Antonnelli, A., Ayala, K., Beckstrom-Starnberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carriaga, K.,  
Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,  
Gupte, J., Haghighi, P., Han, D., Hansen, N., Ho, S.-L., Hu, P.,  
Hurle, B., Idol, J.R., Karlins, E., Kwong, P., Laric, P., Lee-Jin, S.-O.,  
Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masello, C.,  
Maskeri, B., McDowell, J., Pagnigan, C., Pearson, R., Portnoy, M.E.,  
Prasad, A., Reddi-Dugue, N., Schandler, K., Schueler, M.G., Shah, K.,  
Sison, C., Stantipod, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,  
Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 167423)  
Green, E.D.  
Direct Submission  
Submitted (27-JUN-2003) NIH Intramural Sequencing Center, 8717  
Groveport Circle, GaitHERsburg, MD 20877, USA  
3 (bases 1 to 167423)  
Green, E.D.  
Direct Submission  
Submitted (19-JUL-2003) NIH Intramural Sequencing Center, 8717  
Groveport Circle, GaitHERsburg, MD 20877, USA  
On Jul 19, 2003 this sequence version replaced gi:32306558.  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: [nisc\\_zoo@nhgri.nih.gov](mailto:nisc_zoo@nhgri.nih.gov)  
----- Project Information  
Center project name: fcl  
Center clone name: 138F17

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.  
----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 166311 bases at least Q40

Consensus quality: 166956 bases at least Q30  
Consensus quality: 167001 bases at least Q20  
Insert size: 208000; agarose-fp  
Insert size: 167023; sum-of-contigs  
Quality coverage: 9.82x in Q20 bases; agarose-fp  
Quality coverage: 12.23x in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 5 contigs. Gaps between the contigs  
\* are represented as runs of 'N'. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1  
\* 4539: contig of 4539 bp in length  
\* 4540 83205: gap of unknown length  
\* 4640 83205: contig of 78566 bp in length  
\* 83206 83305: gap of unknown length  
\* 83306 155517: contig of 72212 bp in length  
\* 155518 155617: gap of unknown length  
\* 155618 164308: contig of 8691 bp in length  
\* 164309 164408: gap of unknown length  
\* 164409 167423: contig of 3015 bp in length.  
Location/Qualifiers  
1. 167423  
/organism="Macropus eugenii"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9315"  
/clone="ME\_KBa-138F17"  
/clone\_lib="ME\_KBa"  
1. 4539  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"  
4540. 4639  
/estimated\_length=unknown  
4640. 83205  
/note="assembly\_fragment"  
83206. 83305  
/estimated\_length=unknown  
83306. 155517  
/note="assembly\_fragment"  
155518. 155617  
/estimated\_length=unknown  
155618. 164308  
/note="assembly\_fragment"  
164309. 164408  
/estimated\_length=unknown  
164409. 167423  
/note="assembly\_fragment  
clone\_end:77  
vector\_side:right"

ORIGIN

Query Match 91.1%; Score 16.4; DB 14; Length 167423;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18  
|||||

Db 159234 CCCCCCTCTTCGTGATAC 159217

RESULT 38  
BX530055/c 168373 bp DNA linear ROD 25-OCT-2003  
LOCUS Mouse DNA sequence from clone Rp23-462G16 on chromosome X, complete  
DEFINITION  
BX530055  
BX530055  
BX530055.9 GI:37992175  
HTG.



[illegible]

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with 'N's to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

## ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

## ----- Project Information

Center project name: KBHG

Center clone name: CH230-259F3

## ----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 153187 bases at least Q40

Consensus quality: 155728 bases at least Q30

Consensus quality: 157124 bases at least Q20

Estimated insert size: 157936; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

- \* NOTE: Estimated insert size may differ from sequence length
- \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).
- \* NOTE: This is a 'working draft' sequence. It currently
- \* consists of 1 contigs. Gaps between the contigs
- \* are represented as runs of N. The order of the pieces
- \* is believed to be correct as given, however the sizes
- \* of the gaps between them are based on estimates that have
- \* provided by the submitter.
- \* This sequence will be replaced
- \* by the finished sequence as soon as it is available and
- \* the accession number will be preserved.

1 174395: contig of 174395 bp in length.

## Location/Qualifiers

1. 174395

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-259F3"

## misc\_feature

1. 1666

/note="wgs\_end\_extension"

clone\_end:T7"

## misc\_feature

2385. 4003

/note="wgs\_end\_extension"

clone\_end:T7"

## misc\_feature

complement(4619..5461)

/note="clone\_boundary"

clone\_end:T7"

site:

## misc\_feature

end:sequence:B2219160"

## misc\_feature

168236..170264

## misc\_feature

/note="wgs contig"

## misc\_feature

173330..174395

## misc\_feature

/note="wgs\_contig"

## ORIGIN

## Query Match

Best Local Similarity 91.1%; Score 16.4; DB 14; Length 174395;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGGATAC 18

Db 13110 CCCCCCTCTTGGACAC 13127

## RESULT 40

CR792431

178531 bp DNA linear HNG 23-JUN-2005

LOCUS

DEFINITION Danio rerio chromosome 3 clone CH211-285G23, WORKING DRAFT

SEQUENCE, 8 unordered pieces.

## ACCESSION

CR792431

VERSION CR792431.5 GI:68162390

KEYWORDS HNG; HNGS\_PHASE1; HNGS\_DRAFT; HNGS\_FULLTOP.

SOURCE Danio rerio (zebrafish)

## ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Osteichthyes;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 178531)

## REFERENCE

McLaren, S.

Direct Submission

Submitted (17-JUN-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

[zfish-help@sanger.ac.uk](mailto:zfish-help@sanger.ac.uk) Clone requests:[http://www.sanger.ac.uk/Projects/D\\_rerio/fags.shtml#dataight](http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataight)

On Jun 23, 2005 this sequence version replaced gi:53850272.

## COMMENT

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: [zfish-help@sanger.ac.uk](mailto:zfish-help@sanger.ac.uk)

----- Project Information

Center project name: zc285G23

## ----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 174931 bases at least Q40

Consensus quality: 176393 bases at least Q30

Consensus quality: 177177 bases at least Q20

Insert size: 177831; sum-of-contigs

Insert size: 146733; 12.6% error; agarose-gel

Quality coverage: 8.67x in Q20 bases; sum-of-contigs Quality

coverage: 10.93x in Q20 bases; agarose-gel

- \* NOTE: This is a 'working draft' sequence. It currently
- \* consists of 8 contigs. The true order of the pieces
- \* is not known and their order in this sequence record is
- \* arbitrary. Gaps between the contigs are represented as
- \* runs of N, but the exact sizes of the gaps are unknown.
- \* This record will be updated with the finished sequence
- \* as soon as it is available and the accession number will
- \* be preserved.

## FEATURES

## source

1. 178531

/organism="Danio rerio"

/mol\_type="genomic DNA"

/db\_xref="taxon:7955"

/chromosome="13"

/clone="CH211-285G23"

/clone\_11b="CHOR1-211"

1..2317

/note="assembly fragment:00026"

fragment chain:1"

## misc\_feature

2418..12612

/note="assembly fragment:00160"

fragment chain:1"

## misc\_feature

12713..48819

/note="assembly fragment:00599"

misc\_feature fragment\_chain:2"  
48920..60692  
/note="assembly\_fragment:00438  
fragment\_chain:2"  
misc\_feature 60793..71540  
/note="assembly\_fragment:00042"  
71641..157012  
/note="assembly\_fragment:01144.0"  
157113..170004  
/note="assembly\_fragment:01144.1"  
170105..178531  
/note="assembly\_fragment:00285  
clone\_end:T7  
vector\_side:right"  
ORIGIN  
Query Match 91.1%; Score 16.4; DB 14; Length 178531;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCCCCCTCTCTGATAC 18  
Db 90814 CCCCCTTCTTGATAC 90831  
RESULT 41  
CR846078/c 180080 bp DNA linear HTG 08-MAY-2005  
LOCUS Danio rerio clone DKEY-145P15, WORKING DRAFT SEQUENCE, 3 unordered  
DEFINITION pieces.  
ACCESSION CR846078.3 GI:63094510  
VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
KEYWORDS Danio rerio (zebrafish)  
SOURCE Danio rerio  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 180080)  
McLaren,S.  
Direct Submission  
Submitted (07-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On May 8, 2005 this sequence version replaced gi:53748694.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: zf1sh-help@sanger.ac.uk  
----- Project Information  
Center project name: ZK145P15  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 179616 bases at least Q40  
Consensus quality: 179731 bases at least Q30  
Consensus quality: 179800 bases at least Q20  
Insert size: 179880; sum-of-contigs  
Insert size: 185873; 2.1% error; agarose-fp  
Quality coverage: 6.26x in Q20 bases; sum-of-contigs Quality  
coverage: 6.22x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 71753: contig of 71753 bp in length  
\* 71754 71853: gap of 100 bp

\* 71854 143127: contig of 71274 bp in length  
\* 143128 143227: gap of 100 bp  
\* 143228 180080: contig of 3653 bp in length.  
Location/Qualifiers  
source 1..180080  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEY-145P15"  
/clone\_11b="DanioKey"  
1..71753  
/note="assembly\_fragment:01095  
fragment\_chain:1"  
71854..143127  
/note="assembly\_fragment:00374  
fragment\_chain:1"  
143228..180080  
/note="assembly\_fragment:00001.0"  
ORIGIN  
misc\_feature  
misc\_feature  
misc\_feature  
Query Match 91.1%; Score 16.4; DB 14; Length 180080;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCCCCCTCTCTGATAC 18  
Db 178836 CCCCCTTCTTGATAC 178819  
RESULT 42  
CR589883/c 181839 bp DNA linear HTG 09-APR-2005  
LOCUS Danio rerio clone CH211-S1N18, WORKING DRAFT SEQUENCE, 5 unordered  
DEFINITION pieces.  
ACCESSION CR589883  
VERSION CR589883.3 GI:62460894  
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 181839)  
McLaren,S.  
Direct Submission  
Submitted (08-APR-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Apr 9, 2005 this sequence version replaced gi:50724901.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: zf1sh-help@sanger.ac.uk  
----- Project Information  
Center project name: ZC51N18  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 179422 bases at least Q40  
Consensus quality: 179892 bases at least Q30  
Consensus quality: 180245 bases at least Q20  
Insert size: 181439; sum-of-contigs  
Insert size: 184777; 2.8% error; agarose-fp  
Quality coverage: 8.75x in Q20 bases; sum-of-contigs Quality  
coverage: 8.64x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 5 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence

```
* as soon as it is available and the accession number will
* be preserved.
1 18154: contig of 18154 bp in length
* 18155 18254: gap of 100 bp
* 18255 54960: contig of 36706 bp in length
* 54961 55060: gap of 100 bp
* 55061 111004: contig of 55944 bp in length
* 111005 111104: gap of 100 bp
* 111105 173718: contig of 62614 bp in length
* 173719 181839: contig of 100 bp
* 173819 181839: contig of 8021 bp in length.
location/Qualifiers
  source          1..181839
                  /organism="Danio rerio"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:7955"
                  /clone_lib="CH211-51N18"
                  /clone_1lib="CHORI-211"
  1..18154
  /note="assembly fragment:02211"
  fragment_chain:1
  clone_end:T7
  vector_side:left"

misc_feature
  18255..54960
  /note="assembly fragment:00001"
  fragment_chain:1"

misc_feature
  55061..111004
  /note="assembly fragment:00462"
  fragment_chain:1"

misc_feature
  111105..173718
  /note="assembly fragment:01227"
  fragment_chain:1"

misc_feature
  173819..181839
  /note="assembly fragment:02078"
  fragment_chain:1
  clone_end:SP6
  vector_side:right"

ORIGIN

Query Match          91.1%; Score 16.4; DB 14; Length 181839;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18
    |||||
Db 52727 CCCCCTCTTCGTGATAC 52710

RESULT 43
BX569779/c          182464 bp   DNA   linear   VRT 01-APR-2005
LOCUS              Zebrafish DNA sequence from clone CH211-11319 in linkage group 11,
DEFINITION         complete sequence.
ACCESSION          BX569779
VERSION            BX569779.35 GI:62148910
KEYWORDS            HTG.
SOURCE             Danio rerio (zebrafish)
ORGANISM           Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 182464)
Dyer, L.
Direct Submission
Submitted (01-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 1, 2005 this sequence version replaced gi:61969306.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
```

```
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
Zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhifeng Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-11319
is from a CHORI-211 BAC library
VECTOR: pTRABAC2.1.

FEATURES
  source          1..182464
                  /organism="Danio rerio"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:7955"
                  /clone_lib="CH211-11319"
                  /clone_1lib="CHORI-211"
  location/Qualifiers

ORIGIN

Query Match          91.1%; Score 16.4; DB 5; Length 182464;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18
    |||||
Db 93324 CCCCCTCTTCGTGATAC 93307

RESULT 44
AL954390/c          185237 bp   DNA   linear   ROD 09-AUG-2003
LOCUS              Mouse DNA sequence from clone RP23-13H18 on chromosome 4, complete
DEFINITION         sequence.
ACCESSION          AL954390
VERSION            AL954390.19 GI:33569188
KEYWORDS            HTG.
SOURCE             Mus musculus (house mouse)
ORGANISM           Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 185237)
Hopkins, B.
Direct Submission
Submitted (09-AUG-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humney@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Aug 9, 2003 this sequence version replaced gi:33412255.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
```

----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

-----  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, Sw: SWISSPROT, Tr: TREMBL, Wp: WORMPEP, Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-13H18 is from the RPCI-23 Mouse BAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6.  
 Location/Qualifiers  
 1..185237  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-13H18"  
 /clone\_1fb="RPCI-23"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 185237;  
 Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 CCCCCCTCTTGATAC 18  
 |||||  
 Db 140276 CCCCCCTCTTGATAC 140259

RESULT 45  
 CR376821 187216 bp DNA linear VRT 20-JUN-2005  
 LOCUS Zebrafish DNA sequence from clone CH211-67P11 in linkage group 25,  
 DEFINITION complete sequence.  
 ACCESSION CR376821  
 VERSION CR376821 GI:68051166  
 KEYWORDS HTG.  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 187216)  
 Clark, S.  
 REFERENCE Direct Submission  
 TITLE Submitted (18-JUN-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [zfish-help@sanger.ac.uk](mailto:zfish-help@sanger.ac.uk) Clone requests: [http://www.sanger.ac.uk/Projects/D\\_rerio/fags.shtml#database](http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#database)  
 COMMENT On Jun 20, 2005 this sequence version replaced gi:66710823.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>

Contact: [zfish-help@sanger.ac.uk](mailto:zfish-help@sanger.ac.uk)

-----  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, Sw: SWISSPROT, Tr: TREMBL, Wp: WORMPEP, Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) Clone-derived Zebrafish pBC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats, where this is found the longest good quality representation will be submitted.

## FEATURES

## source

Location/Qualifiers  
 1..187216  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /chromosome="25"  
 /clone="CH211-67P11"  
 /clone\_1fb="CHORI-211"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 5; Length 187216;  
 Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 CCCCCCTCTTGATAC 18  
 |||||  
 Db 48792 CCCCCCTCTTGATAC 48809

RESULT 46  
 AC155840/C 188181 bp DNA linear ROD 30-APR-2005  
 LOCUS Mus musculus 6 BAC RP24-223C8 (Roswell Park Cancer Institute  
 DEFINITION (57BL/6J Male) Mouse BAC Library) complete sequence.  
 ACCESSION AC155840 AC119891  
 VERSION AC155840 GI:62988423  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.  
 1 (bases 1 to 188181)  
 Muzny, D., Adams, C., Agbai, I. O., Allen, C., Albrooks, S., Archer, P., Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R., Berarducci, A., Biswal, K., Blyth, P., Bonham, H., Bunay, C., Burch, P., Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, T., Chacko, J., Chahrouh, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Dai, W., Davila, M. L., Davis, C., Day-Carroll, L., De Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K., Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Bugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gaekin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,

[illegible]

```

misc_feature      /clone="RP24-223C8"
                  complement(1..6647)
                  /note="overlaps bases 1..66471 of clone AC130220"
misc_feature      /function="clone overlap"
                  15073..15215
                  /function="single clone coverage"
misc_feature      50015..50026
                  /function="single clone coverage"
misc_feature      83586..83808
                  /function="pcr product sequence only"
misc_feature      101036
                  /function="bacterial transposon excised"
misc_feature      176064..176793
                  /function="unresolved tandem repeat"
misc_feature      183650..188181
                  /note="overlaps bases 1..4532 of clone AC153997"
                  /function="clone overlap"

ORIGIN
Query Match      91.1%; Score 16.4; DB 9; Length 188181;
Best Local Similarity 94.4%; Pred. No.5.e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy               1 CCCCCCTCTTCGTGATAC 18
                |||||
Db              171180 CCCCCTCTCTGTGTGATAC 171163

RESULT 47
LOCUS            AL591598             191520 bp    DNA          linear   ROD 24-JUN-2002
DEFINITION       Mouse sequence from clone RP23-272C14 on chromosome 2, complete
VERSION          AL591598
KEYWORDS         HTG.
SOURCE           Mus musculus (house mouse)
ORGANISM         Mus musculus
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Embryophyta; Eumetazoa; Euarchontoglires; Glires; Rodentia;
                 Sciurgnathi; Muroidae; Murinae; Mus.
                 1 (bases 1 to 191520)
                 Leongamornlert,D.
                 Direct Submission
                 Submitted (22-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
                 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
                 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
                 On Apr 29, 2002 this sequence version replaced gi:19699444.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by association digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; SW:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPI21-Mouse PAC Library

```

constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACe3.6.  
Location/Qualifiers

1.191520  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="2"  
/clone="RP23-272C14"  
/clone\_1ib="RPc1-23"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 191520;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCCCCCTCTTCGATAC 18  
Db 111972 CCTCCCTCTTCGATAC 111955

RESULT 48  
AC153716/c  
LOCUS  
DEFINITION  
Bos taurus clone CH240-21E9, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 19  
unordered pieces.  
AC153716  
AC153716.2 GI:68265877  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
SOURCE  
Bos taurus (cow)  
ORGANISM  
Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
Pecora; Bovidae; Bovinae; Bos.  
1 (bases 1 to 192040)  
Muzny,D,Marle,,Metzker,M,Lee,,Abramson,S,,Adams,C,,Alder,J.,  
Allen,C,,Allen,H,,Alsbrechts,S,,Amin,A,,Anguiano,D.,  
Anyalebechi,V.,Aoyagi,A.,Ayodeji,M.,Bacca,E.,Baden,H.,  
Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,  
Biswal,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,  
Bryant,N.,Buhay,C.,Burich,P.,Burrell,K.,Calderson,B.,  
Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Center,A.,  
Chacho,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,  
Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,  
Davila,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dederich,D.,  
Deigado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Diya,K.,  
Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duvall,B.,Baves,K.,  
Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,  
Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,  
Fraser,C.M.,Gabriel,A.,Ganta,R.,Garcia,A.,Garner,T.,Gazda,W.,  
Gibberson,G.,Gier,K.,Gill,R.,Grady,M.,Guerra,M.,Guevara,M.,  
Gunnarsson,P.,Haaland,W.,Hamil,C.,Hamilton,C.,Hamilton,K.,  
Harvey,Y.,Havlik,P.,Hawes,A.,Henderson,N.,Hernandez,J.,  
Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgeson,A.,Hogues,M.,  
Hollins,B.,Howells,S.,Huliy,S.,Hume,J.,Idlebirt,D.,Jackson,A.,  
Jackson,B.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolivet,A.,  
Karpach,S.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Kovar,C.,  
Kowis,C.,Kraft,C.L.,Lebow,H.,Levan,J.,Lewis,J.,Li,Z.,Liu,J.,  
Liu,J.,Liu,W.,Liu,Y.,London,P.,Longacre,S.,Lopez,D.,  
Lorenshaw,L.,Louisgied,H.,Lopez,R.J.,Lu,X.,Ma,J.,  
Maheshwari,M.,Mahindartine,M.,Mahmoud,M.,Malloy,K.,Mangum,A.,  
Mangum,B.,Mapa,P.,Martin,K.,Martin,R.,Martinez,E.,  
Mawhinney,S.,McLeod,M.P.,McNeill,T.Z.,Meenen,E.,  
Milosavljevic,A.,Miner,G.,Mintja,B.,Montemayor,J.,Moore,S.,  
Morgan,M.,Morris,K.,Morris,S.,Munidas,M.,Murphy,M.,Nair,L.,  
Nankervis,C.,Neal,D.,Newton,N.,Nguyen,N.,Norris,S.,  
Nsoekiemeh,O.,Okunolu,G.,Olariunpasegon,A.,Pal,S.,Parks,K.,  
Pasternak,S.,Paul,H.,Perez,A.,Perez,L.,Pfannkuch,C.,  
Plopper,F.,Polinder,A.,Popovic,D.,Primus,E.,Pu,L.-L.,  
Pauz,M.,Quiroz,J.,Rachin,B.,Reeves,K.,Regier,M.A.,Reigh,R.,  
Reilly,B.,Reilly,M.,Ren,Y.,Reuter,M.,Richard,S.,Rigga,F.,  
Rives,C.,Rodkey,T.,Rojas,A.,Rose,M.,Rose,R.,Ruiz,S.J.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Sanders,W.,Savery,G.,Scherer,S.,Scott,G.,Shateman,S.,Shen,H.,  
Shetty,J.,Shwartzbein,A.,Sison,I.,Sitter,C.D.,Smaj,D.,  
Sneed,A.,Sodergren,B.,Song,X.-Z.,Sorelle,R.,Sosa,J.,  
Steinle,M.,Strong,R.,Sutton,A.,Svatek,A.,Tabot,P.,Taylor,C.,  
Taylor,T.,Thomas,N.,Thomas,S.,Tingey,A.,Trejos,Z.,Uman,K.,  
Valas,R.,Vera,V.,Villasana,D.,Waldron,L.,Walker,B.,Wang,J.,  
Wang,Q.,Wang,S.,Warren,J.,Warren,R.,Wei,X.,White,F.,  
Williams,G.,Wilson,R.,Wiczek,R.,Wooden,H.,Worley,K.,  
Wright,D.,Wright,R.,Wu,J.,Yakub,S.,Yen,J.,Yoon,L.,Yoon,V.,  
Yu,F.,Zhang,J.,Zhou,J.,Zhou,X.,Zhao,S.,Zhou,D.,von  
Niederhausern,A.,Weiss,R.,Smith,D.R.,Holt,R.A.,Smith,H.O.,  
Weinstock,G. and Gibbs,R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 192040)  
Worley,K.C.  
Direct Submission  
Submitted (15-DEC-2004) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 192040)  
Cow Genome Sequencing Consortium.  
Direct Submission  
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jun 28, 2005 this sequence version replaced gi:56606215.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center Project name: FBIM  
Center clone name: CH240-21E9  
----- Summary Statistics  
Assembly program: Atlas 3.0  
Consensus quality: 182840 bases at least Q40  
Consensus quality: 184826 bases at least Q30  
Consensus quality: 186385 bases at least Q20  
Estimated insert size: 183715; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 19 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
1 1222: contig of 1222 bp in length  
1223 1704: gap of 482 bp  
1705 13754: contig of 12050 bp in length  
13755 13804: gap of 50 bp  
13805 50265: contig of 36461 bp in length  
50266 50315: gap of 50 bp  
50316 58874: contig of 8559 bp in length  
58875 59802: gap of 928 bp

```
* 59803 78202: contig of 18400 bp in length
* 78203 78252: gap of 50 bp
* 78253 79290: contig of 1038 bp in length
* 79291 79589: gap of 299 bp
* 79590 90739: contig of 11150 bp in length
* 90740 90789: gap of 50 bp
* 90790 111950: contig of 21161 bp in length
* 111951 112000: gap of 50 bp
* 112001 151644: contig of 39644 bp in length
* 151645 151694: gap of 50 bp
* 151695 152984: contig of 1290 bp in length
* 152985 153806: gap of 823 bp
* 153807 155975: contig of 2169 bp in length
* 155976 156025: gap of 50 bp
* 156026 178376: contig of 22351 bp in length
* 178377 178827: gap of 451 bp
* 178828 180376: contig of 1449 bp in length
* 180377 180376: gap of unknown length
* 180377 181394: contig of 1018 bp in length
* 181395 181494: gap of unknown length
* 181495 183187: contig of 1693 bp in length
* 183188 183287: gap of unknown length
* 183288 184305: contig of 1018 bp in length
* 184306 184405: gap of unknown length
* 184406 186952: contig of 2547 bp in length
* 186953 187052: gap of unknown length
* 187053 189112: contig of 2060 bp in length
* 189113 189212: gap of unknown length
* 189213 192040: contig of 2828 bp in length.
```

FEATURES  
SOURCE

```
1..192040
location=Qualifiers
  /organism="Bos taurus"
  /mol_type="genomic DNA"
  /db_xref="taxon:9913"
  /clone="CH240-21E9"
1223..1704
  /estimated_length=482
13755..13804
  /estimated_length=50
50266..50315
  /estimated_length=50
58875..59802
  /estimated_length=928
78203..78252
  /estimated_length=50
79291..79589
  /estimated_length=299
90740..90789
  /estimated_length=50
111951..112000
  /estimated_length=50
151645..151694
  /estimated_length=50
152985..153806
  /estimated_length=822
155976..156025
  /estimated_length=50
178377..178827
  /estimated_length=451
180277..180376
  /estimated_length=unknown
181395..181494
  /estimated_length=unknown
183188..183287
  /estimated_length=unknown
184306..184405
  /estimated_length=unknown
186953..187052
  /estimated_length=unknown
189113..189212
  /estimated_length=unknown
```

ORIGIN

```
Query Match          91.1%; Score 16.4; DB 14; Length 192040;
Beet Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1  CCCCCCTCTTCGTGATAC 18
          |||
Db       120679  CCTCCTCTTCGTGATAC 120662
```

## RESULT 49

BX322665

## DEFINITION

BX322665 192437 bp DNA linear VRT 17-APR-2005  
Zebrafish DNA sequence from clone DKEY-183N20 in linkage group 20  
Contains two novel genes, the cyp2j2a gene for cytochrome P450  
family 2 subfamily J polypeptide 2 A, the cyp2j2b gene for  
cytochrome P450, family 2 subfamily J polypeptide 2 B, the cyp2v2  
gene for cytochrome P450 family 2 subfamily V polypeptide 2, five  
genes for novel proteins similar to cytochrome P450 family 2  
subfamily J, the gene for a novel protein similar to cytochrome  
P450 family 2 subfamily J (zgc:55731), the cyp2j1a gene for  
cytochrome P450 subfamily IIU (arachidonic acid epoxidase)  
polypeptide A, the gene for a novel protein similar to vertebrate  
hook homolog 1 (Drosophila) (HOOK1), the gene for a novel zinc  
finger protein, the gene for a novel protein similar to vertebrate  
phosphoribosylformylglycinamide synthase (PGR amidotransferase)  
(PPAS), the gene for a novel protein similar to vertebrate  
mitochondrial ribosomal protein L19 (MRPL19) and a novel  
pseudogene, complete sequence.

## ACCESSION

BX322665

## VERSION

HTG; cyp2j2a; cyp2j2b; cyp2j1a; cyp2v2; zgc:55731.

## KEYWORDS

Danio rerio (zebrafish)

## SOURCE

Danio rerio

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 192437)

Tromans,A.

Direct Submission

Submitted (02-Dec-2004) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

fish-help@sanger.ac.uk Clone requests:

http://www.sanger.ac.uk/Projects/D\_rerio/fags.shtml#dataight

On Nov 20, 2003 this sequence version replaced gi:31335455.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: zfish-help@sanger.ac.uk

-----

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WormBEP; Information on the WormBEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormBEP](http://www.sanger.ac.uk/Projects/C_elegans/wormBEP) DKEY-183N20 is from a Zebrafish BAC library

VECTOR: pindigobAC-5  
Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

Any regions longer than 1kb tagged as misc-feature 'unseq' are part of a tandem repeat of more than 10kb in length where it has not been possible to anchor the base differences between repeat copies. The region has been built up based on the repeat element to match the total size of repeat indicated by restriction digest,



but repeat copies may not be in the correct order and the usual  
finishing criteria may not apply.

FEATURES  
Location/Qualifiers

source

1..192437

/organism="Dario reio"

/mol\_type="genomic DNA"

/db\_xref="taxon:7955"

/chromosome="20"

/clone\_name="DKEY-183N20"

/clone\_lib="Dantolkey"

8007..8012

polyA\_signal

polyA\_site

polyA\_site

polyA\_signal

gene

mRNA

complement(14530)

complement(14551..14566)

complement(14557..14562)

complement(join(22815..23427,23548..23729,25153..25287,25398..25516,27163..27274,27456..27547))

/locus\_tag="DKEY-183N20.17-001"

complement(join(22815..23427,23548..23729,25153..25287,25398..25516,27163..27274,27456..27547))

/locus\_tag="DKEY-183N20.17-001"

/product="novel protein similar to vertebrate

mitochondrial ribosomal protein L19 (MRPL19)"

/note="match: ESTs: Em:AI588392.1 Em:BI67355.1

Em:BM182601.1 Em:BM774932.1 Em:CK029938.1 Em:CK354603.1"

complement(join(22815..23427,23548..23729,25153..25287,25398..25516,27163..27274,27456..27547))

/locus\_tag="DKEY-183N20.17-002"

complement(join(22815..23427,23548..23729,25153..25287,25398..25516,27163..27274,27456..27547))

/locus\_tag="DKEY-183N20.17-002"

/product="novel protein similar to vertebrate

mitochondrial ribosomal protein L19 (MRPL19)"

/note="match: ESTs: Em:CK029254.1"

complement(join(23206..23427,23548..23729,25153..25287,25398..25516,27163..27274,27456..27547))

/locus\_tag="DKEY-183N20.17-002"

/standard\_name="OTDARP0000005362"

/codon\_start=1

/product="novel protein similar to vertebrate

mitochondrial ribosomal protein L19 (MRPL19)"

/protein\_id="CAH69005.1"

/db\_xref="GI:55251116"

/db\_xref="InterPro:IPR001857"

/db\_xref="InterPro:IPR008279"

/db\_xref="UniProt/TREMBL:Q5T723"

/translation="MAACRRARERMEGLRLNIPOTERFLTSVNRHASDGRSKFI

PPTKVPPTDSQSEASVRRVLSPEFIPRROTQIKFYIERKMIQRKVLQIPFV

GSILAVTMDPYSAGNLRFVIGICTORSKGLATFLLRNVIDGQELCYELYSFPMK

IVYLKLEKRLDDNIMLYRLDALPEYSTDPMQPVHVELTKDIPNPLKMKKPEPMK

RWERPKFDIKGIRDLVLPBQMAKQKGEPMREYMLKEYDSSLEKKILLEVDSN

LRR"

complement(join(23206..23427,23548..23729,25153..25287,25398..25516,27163..27274,27456..27547))

/locus\_tag="DKEY-183N20.17-002"

/standard\_name="OTDARP0000005361"

/note="match: proteins: Sw:P49406 Sw:Q9D338 Tr:AAH30144

Tr:BAC34357"

/codon\_start=1

/product="novel protein similar to vertebrate

mitochondrial ribosomal protein L19 (MRPL19)"

/protein\_id="CAH69004.1"

/db\_xref="GI:55251115"

/db\_xref="GOA:Q6DC47"

/db\_xref="InterPro:IPR001857"

/db\_xref="UniProt/TREMBL:Q6DC47"

/translation="MAACRRARERMEGLRLNIPOTERFLTSVNRHASDGRSKFI

PPTKVPPTDSQSEASVRRVLSPEFIPRROTQIKFYIERKMIQRKVLQIPFV

GSILAVTMDPYSAGNLRFVIGICTORSKGLATFLLRNVIDGQELCYELYSFPMK

RWERPKFDIKGIRDLVLPBQMAKQKGEPMREYMLKEYDSSLEKKILLEVDSN

polyA\_site

polyA\_signal

gene

mRNA

complement(23268)

/locus\_tag="DKEY-183N20.17-002"

complement(23281..23286)

/locus\_tag="DKEY-183N20.17-002"

28127..55004

/locus\_tag="DKEY-183N20.16-001"

join(28127..28205,28789..28916,30456..30588,30666..30771,30858..31056,32535..32640,33020..33160,33246..33370,34141..34269,34435..34566,34645..34773,35847..35920,34125..41281,41679..41740,41827..41988,42072..42176,42841..42972,46137..46367,46866..47046,48340..48423,48510..48665,48830..48949,50684..50829,51131..51338,51423..51514,52338..52452,52675..52821,52999..53139,54123..55004)

/locus\_tag="DKEY-183N20.16-001"

/product="novel protein similar to vertebrate

phosphoribosylformylglycinamide synthase (FGAR

amidotransferase) (PFAS)"

/note="match: ESTs: Em:AI584881.1 Em:AI598852.1

Em:AI558857.1 Em:BG737723.1 Em:BI890093.1 Em:CA470200.1

Em:CD281609.1 Em:CD281610.1 Em:CK022922.1"

join(28793..28916,30456..30588,30666..30771,30858..31056,32535..32640,33020..33160,33246..33370,34141..34281,34435..34566,34645..34773,35847..35920,41125..41281,41679..41740,41827..41988,42072..42176,42841..42972,46137..46367,46866..47046,48340..48423,48510..48665,48830..48949,50684..50829,51133..51338,51423..51514,52338..52452,52675..52821,52999..53139,54123..55004)

/locus\_tag="DKEY-183N20.16-001"

/standard\_name="OTDARP0000005369"

/note="match: proteins: Sw:O15067 Sw:Q9WBD3 Tr:AAH82811

Tr:Q7PYT0 Tr:Q84XV8 Tr:Q8L4G7"

/codon\_start=1

/product="novel protein similar to vertebrate

phosphoribosylformylglycinamide synthase (FGAR

amidotransferase) (PFAS)"

/protein\_id="CAH69006.1"

/db\_xref="GI:55251117"

/db\_xref="GOA:Q5T722"

/db\_xref="InterPro:IPR007028"

/db\_xref="InterPro:IPR010073"

/db\_xref="InterPro:IPR010918"

/db\_xref="UniProt/TREMBL:Q5T722"

/translation="MPVPRRYRREBGERAARIQAVLPVITTELQVNEVDGPD

SLSVAQDILRLMFSPPYSVLSEPTLRLKHAQRLVEIPRANFSNMTNVSICO

SAGISQVTRVELSRRLIKPQECQKVMQGEESLISCLYDSTECTVIAQPTTSFV

DIRPDVEVDIILGKRAALEKANDEIGLAFDWDLDLYTALFOKVRNPTSVCEFL

AGNSHSRHMFRGRNVLDGKQKTELFLINGTOOHSQNNVIFKCDNSGIGKME

LRCHYPTNPAQASDYESRDTTRHVIFPAETHNPTGVAFPGSANTGGRIRVOVAG

KGGHVLATGTCFPGNHLHGFPLPMBEGBEPPSSAPRLOVATASDASGNKFK

GEEVLGFAFSFMRLANGERRRWIKPMSGGVLSIEDPVRKQDAPEPMVYKICG

PVIRIGVGGGAASVOVQDNGSRARDGAFDQDAEQKQNALRCLERVGNPLFC

SHDOAGNGNVLKELSEBPAGAVITYEKKREDPTLSVLELGAEOESNALLRKS

DRSEFLRVCORKECPVDPFVCKITGDGTVLDGRLKONVLEBGRNVDLELMDVLK

MPQKPTFLERSVSLQPLTPALGSYLPALERLRLPAYSXKRAKLTNKVSTGLVA

QOOCVGLPLRLADVAVALSPSLQCATATICEOPTKGLSSKAGARMAVGATLTL

VFPVRSALKQVKSQMMWMAKRGESACATACQACQCEVMGOLGAVDGCKSLSNA

ARVSGETVAKPSGLVSVVAPCDIATVPTLIDRLNPEKGVLLVVPASAKYRIGSA

LAOCFQLGDCSPMDQPDLSACFNTQTLIDRLTLTAGHDVSGLLISCLLEMAFA

GNYGIYDPLBEGVDWEALFSELSLWVSECENASVCAVYDADALGLRIGSTTG

FGPDPAKRVSLCREVYNERLPTLRALWVSTSELELQANPLCVQEEBQGLARTOP

YLUKTPDSQTPILIKELATKARAVAVREBSNGDRMSMSLPMAGFEVMDMODIC

SGSFTLDPFRAVVFGEFSDYADVGSKGAATVTFNNRABEERFKKEDTSLGV

CNGCOLLALIGWVGERDGSVDYTLTNKSGRFSRVSGLIPLSPAIMLKGMGSAL

GWWVAHEGMLQFRSPEAQOKLIGSSILAPRYVDSGNPTEIYIPINPNSQGVAGIC

Query Match

Similarity 91.1%

Score 16.4

DB 5

Length 192437

Beet Local

Matches 17

Conservative 0

Mismatches 1

Indels 0

Gaps 0

1

CCCCCTCTCTGATAC 18

CCCCCTCTCTGATAC 181613



```

gap /note="assembly_fragment"
    4288.4387
misc_feature /estimated_length=100
    4388.11362
gap /note="assembly_fragment"
    11363.11462
misc_feature /estimated_length=100
    11463.22942
gap /note="assembly_fragment"
    22943.23042
misc_feature /estimated_length=100
    23043.41668
gap /note="assembly_fragment"
    41669.41768
misc_feature /estimated_length=100
    41769.90764
gap /note="assembly_fragment"
    90765.90864
misc_feature /estimated_length=100
    90865.110835
gap /note="assembly_fragment"
    110836.110935
misc_feature /estimated_length=100
    110936.139954
gap /note="assembly_fragment"
    139955.140054
misc_feature /estimated_length=100
    140055.181300
gap /note="assembly_fragment"
    181301.181400
misc_feature /estimated_length=100
    181401.195220
gap /note="assembly_fragment
    clone_end.T7
    vector_side:right"

```

## ORIGIN

```

Query Match      91.1%; Score 16.4; DB 14; Length 195220;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCCCCCTTCTGTGATAC 18
Db 109600 CCCCCCTTCTGTGATAC 109617

```

```

RESULT 51
CR759830 196033 bp DNA linear VRT 06-FEB-2005
LOCUS zebrafish DNA sequence from clone DKEXP-79F12 in linkage group 23,
DEFINITION complete sequence.
ACCESSION CR759830
VERSION CR759830.8 GI:58651984
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 196033)
Clark.S.
REFERENCE Direct Submission
AUTHORS Submitted (06-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambidgehire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Feb 4, 2005 this sequence version replaced gi:58037040.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.

```

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats, where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see [http://www.sanger.ac.uk/Projects/D\\_rerio/fishmark.shtml](http://www.sanger.ac.uk/Projects/D_rerio/fishmark.shtml) DKEXP-79F12 is from a Zebrafish BAC library

VECTOR: pindigobac-5.

## FEATURES

```

source
    1.196033
    /organism="Danio rerio"
    /mol_type="genomic DNA"
    /db_xref="taxon:7955"
    /clone="DKEXP-79F12"
    /clone_1ib="DanioKeyPilot"

```

## ORIGIN

```

Query Match      91.1%; Score 16.4; DB 5; Length 196033;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCCCCCTTCTGTGATAC 18
Db 107334 CCCCCCTTCTGTGATAC 107317

```

```

RESULT 52
AC117089 197981 bp DNA linear HTG 15-NOV-2002
LOCUS Rattus norvegicus clone CH230-119G9, WORKING DRAFT SEQUENCE, 2
DEFINITION unordered pieces.
ACCESSION AC117089
VERSION AC117089.4 GI:25008560
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 197981)
Muzny,D.,Marle, Metzker,M.,Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooke,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnesstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,B.,
Cardenas,V., Carter,K., Cavazos,I., Cesari,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davis,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Driper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

```

Egan, A., Escotro, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flaggs, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregergis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowib, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, L., Louie, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okunolu, G., Olarinmoye, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, T., Sitter, C.D., Smaiz, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soosa, J., Steinhilber, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanol, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, K., White, F., Williams, G., Willson, R., Wlezyk, R., Wood, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, Y., Yoon, V., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished  
2 (bases 1 to 197981)  
Worley, K.C.  
Direct Submission  
Submitted (08-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 197981)  
Rat Genome Sequencing Consortium.  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:23194925.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center Project name: GTFP  
Center clone name: CH230-119G9  
----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 189549 bases at least Q40  
Consensus quality: 190935 bases at least Q20  
Consensus quality: 192081 bases at least Q20  
Estimated insert size: 196342; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\*\*\*\*\*  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 196422: contig of 196422 bp in length  
\* 196423 196522: gap of unknown length  
\* 196523 197981: contig of 1459 bp in length.  
Location/Qualifiers  
1. 197981  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-119G9"  
194929..195819  
/note="clone\_boundary  
site:  
end sequence:BH315129"  
196423..196522  
/estimated\_length=unknown

gap

ORIGIN

Query Match 91.1%; Score 16.4; DB 14; Length 197981;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGATAC 18  
|||||  
Db 191745 CCCCCCTCTTCGATAC 191762

RESULT 53  
AC160068/c  
LOCUS  
DEFINITION  
Bos taurus clone CH240-87H4, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 24  
unordered pieces.  
ACCESSION  
AC160068 2 GI:68301368  
VERSION  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
KEYWORDS  
Bos taurus (cow)  
SOURCE  
ORGANISM  
Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
Pecora; Bovidae; Bovinae; Bos.  
1 (bases 1 to 197993)  
Muzny, D., Marie, Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amn, A., Angilano, D., Anyalebech, V., Ayogaji, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Detamo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotro, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flaggs, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL
Georgievski, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Huliyil, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowik, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzulewa, L., Louisedge, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmood, M., Malloy, K., Mangun, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mashiney, S., Mcleod, M. P., McNeill, T. Z., Meenen, E., Mlosoevljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelenoh, O., Okunodu, G., Olampunsaogun, A., Pal, S., Parks, K., Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Piopper, F., Poidexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Rella, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Tjio, J. Z., Uman, K., Valas, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleciyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.	Submitted (01-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 3 (bases 1 to 197999) Cow Genome Sequencing Consortium. Direct Submission Submitted (01-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	Submitted (18-APR-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 3 (bases 1 to 197999) Cow Genome Sequencing Consortium. Direct Submission Submitted (01-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	Submitted (18-APR-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 3 (bases 1 to 197999) Cow Genome Sequencing Consortium. Direct Submission Submitted (01-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	Submitted (18-APR-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 3 (bases 1 to 197999) Cow Genome Sequencing Consortium. Direct Submission Submitted (01-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

```

Consensus quality: 19349 bases at least Q20
Estimated insert size: 195717, sum-of-contigs estimation
Quality coverage: 6x in Q20 bases, sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 7654: contig of 7654 bp in length
* 7655 7704: gap of 50 bp
* 7705 14857: contig of 7153 bp in length
* 14858 14957: gap of unknown length
* 14958 19375: contig of 4418 bp in length
* 19376 19425: gap of 50 bp
* 19426 22917: contig of 3492 bp in length
* 22918 23244: gap of 327 bp
* 23245 53391: contig of 30147 bp in length
* 53392 53441: gap of 50 bp
* 53442 80304: contig of 26863 bp in length
* 80305 80354: gap of 50 bp
* 80355 84197: contig of 3843 bp in length
* 84198 84247: gap of 50 bp
* 84248 89173: contig of 4926 bp in length
* 89174 89223: gap of 50 bp
* 89224 95022: contig of 5799 bp in length
* 95023 95072: gap of 50 bp
* 95073 107040: contig of 11968 bp in length
* 107041 107090: gap of 50 bp
* 107091 112069: contig of 4979 bp in length
* 112070 112119: gap of 50 bp
* 112120 123496: contig of 11377 bp in length
* 123497 123546: gap of 50 bp
* 123547 127750: contig of 4204 bp in length
* 127751 127800: gap of 50 bp
* 127801 131244: contig of 3444 bp in length
* 131245 131294: gap of 50 bp
* 131295 143535: contig of 12242 bp in length
* 143536 143586: gap of 50 bp
* 143587 155123: contig of 11537 bp in length
* 155124 155173: gap of 50 bp
* 155174 169019: contig of 13846 bp in length
* 169020 169069: gap of 50 bp
* 169070 184782: contig of 15713 bp in length
* 184783 184882: gap of unknown length
* 184883 184895: contig of 1013 bp in length
* 184896 185955: gap of unknown length
* 185956 187135: contig of 1140 bp in length
* 187136 187235: gap of unknown length
* 187236 188361: contig of 1126 bp in length
* 188362 188461: gap of unknown length
* 188462 190596: contig of 2135 bp in length
* 190597 190696: gap of unknown length
* 190697 195789: contig of 5093 bp in length
* 195790 195889: gap of unknown length
* 195890 197999: contig of 2110 bp in length.
*
* Location/Qualifiers
* 1..197999
* /organism="Bos taurus"
* /mol_type="genomic DNA"
* /db_xref="taxon:9913"
* /clone="CR240-87H4"
*
* 7655..7704
* /estimated_length=50
* 14858..14957
* /estimated_length=unknown
* 19376..19425
* /estimated_length=50

```

```

gap      22918..23244
          /estimated_length=327
gap      53392..53441
          /estimated_length=50
gap      80305..80354
          /estimated_length=50
gap      84198..84247
          /estimated_length=50
gap      89174..89223
          /estimated_length=50
gap      95023..95072
          /estimated_length=50
gap      107041..107090
          /estimated_length=50
gap      112070..112119
          /estimated_length=50
gap      123497..123546
          /estimated_length=50
gap      127751..127800
          /estimated_length=50
gap      131245..131294
          /estimated_length=50
gap      143537..143586
          /estimated_length=50
gap      155124..155173
          /estimated_length=50
gap      169020..169069
          /estimated_length=50

```

```

Query Match      91.1%; Score 16.4; DB 14; Length 197999;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 CCCCCCTCTTCGTGATAC 18
          |||||
Db      53934 CACCCCTCTTCGTGATAC 53917

```

```

RESULT 54
LOCUS      AL935173                205213 bp      DNA      linear      VRT 17-APR-2003
DEFINITION Zebrafish DNA sequence from clone CH211-196D9, complete sequence.
ACCESSION  AL935173
VERSION     AL935173.9      GI:30024454
KEYWORDS
SOURCE      HTG:
            Danio rerio (zebrafish)
            Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
            1 (bases 1 to 205213)
            Pelan,S.
REFERENCE   Direct Submission
            Submitted (17-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            On Apr 18, 2003 this sequence version replaced gi:2988588.
COMMENT     ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)

beginning 'Dr' were identified by the Recon repeat discovery system (Zhifeng Bao and Sean Eddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see [http://www/Projects/D\\_rerio/fishmask.shtml](http://www/Projects/D_rerio/fishmask.shtml)

CH211-196D9 is from a CHORI-211 BAC library

VECTOR: pPRABAC2.1.

```

FEATURES
    source
        location/Qualifiers
            1..205213
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /clone="CH211-196D9"
            /clone_lib="CHORI-211"

```

```

Query Match      91.1%; Score 16.4; DB 5; Length 205213;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 CCCCCCTCTTCGTGATAC 18
          |||||
Db      90300 CCCCCCTCTTCGTGATAC 90317

```

```

RESULT 55
LOCUS      AC162927                213948 bp      DNA      linear      ROD 28-JUN-2005
DEFINITION Mus musculus BAC clone RP23-269E12 from chromosome 9, complete
sequence.
ACCESSION  AC162927
VERSION     AC162927.3      GI:68268258
KEYWORDS
SOURCE      Mus musculus (house mouse)
            Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 213948)
            Schmidt,H. and Levy,A.
            The sequence of Mus musculus BAC clone RP23-269E12
            Unpublished (2001)
            2 (bases 1 to 213948)
            Wilson,R.K.
REFERENCE   Direct Submission
            Submitted (02-JUN-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
            3 (bases 1 to 213948)
            Wilson,R.K.
REFERENCE   Direct Submission
            Submitted (09-JUN-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
            4 (bases 1 to 213948)
            Wilson,R.K.
REFERENCE   Submitted (28-JUN-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            On Jun 28, 2005 this sequence version replaced gi:67078759.
COMMENT     ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: submissions@wustl.edu
            Summary Statistics
            Center project name: M_BA0269E12

```

## NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e. phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence. The assembly was confirmed by restriction digest. This finishing standard has slightly changed from the previous Human standard. Specifically, standards for regions of low sequence complexity (such as dinucleotide repeats and small unit tandem repeats) have been relaxed. These regions are very prevalent in the mouse genome, and the return on extended finishing efforts is minimal.

If a sequence meets the criteria of the above statement, it needs no comments or tags. If the criteria are not met, such as ambiguous bases, then the region is duly annotated.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

## SOURCE INFORMATION:

The BAC library has been constructed by Kazuhiro Oseegawa and Minako Tatenio in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

This sequence is the entire insert of the clone.

## Location/Qualifiers

```
1. 213948
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/clone="RP23-269E12"
/clone_1lb="RRC1-23"
193158..193333
/note="Unresolved simple sequence repeat."
207517..207842
/note="Unresolved simple sequence repeat."
```

## ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 213948;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Oy 1 CCCCCCTTCTGATAC 18
    |||||
Db 158917 CCCCCCTTCTGATAC 158934
```

## RESULT 56

## HTCRBCH9/c

LOCUS HTCRBCH9 216293 bp DNA linear PRI 30-OCT-2002  
DEFINITION Homo sapiens chromosome 9 duplication of the T cell receptor beta locus and tryptsinogen gene families.

## ACCESSION

AF029308  
AF029308.1 GI:2564750

## KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

## REFERENCE

1 (bases 1 to 216293)  
Rowen,L., Traak,B., Boyesen,C., Qin,S., Wang,K., Ahearn,M.E. and Hood,L.

## TITLE

Sequence of a large duplication from human chromosome 7 to chromosome 9 containing a portion of the T cell receptor beta locus and tryptsinogen locus

## JOURNAL

Unpublished  
2 (bases 1 to 216293)

## AUTHORS

Rowen,L., Traak,B., Boyesen,C., Qin,S., Wang,K., Ahearn,M.E. and Hood,L.

## TITLE

Direct Submission  
Submitted (09-OCT-1997) Department of Molecular Biotechnology, Box 357730 University of Washington, Seattle, Washington 98195, USA

## COMMENT

Sequencing methodology: high redundancy shotgun, interspersed repeats were identified with RepeatMasker (available from <http://ftp.genome.washington.edu/RM/RepeatMasker.html>) Simple sequence repeats were identified with sputnik (available from <http://serc.mbl.washington.edu/~chrisa/software/sputnik.html>).

## FEATURES

## source

## Location/Qualifiers

```
1. 216293
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
1..37851
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
```

## source

## Location/Qualifiers

```
/clone="cosmid X224"
/clone_1lb="Kai Wang's ATCC1475"
37848..49224
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
```

## source

## Location/Qualifiers

```
/clone="cosmid X91"
/clone_1lb="Kai Wang's ATCC1475"
39045..50577
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
/clone="cosmid G2724"
/clone_1lb="G27"
40575..61633
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
```

## source

## Location/Qualifiers

```
/clone="cosmid 112"
/clone_1lb="CGM-1 YAC V10C11"
/clone_1lb="CGM-1 YAC V10C11"
47024..87062
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
/clone="cosmid 3A10"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
75308..116845
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
```

## source

## Location/Qualifiers

```
/clone="cosmid 3A10"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
75308..116845
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
/clone="cosmid 3A10"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
75308..116845
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
```

## source

## Location/Qualifiers

```
/clone="cosmid 3A10"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
75308..116845
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
/clone="cosmid 3A10"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
75308..116845
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
```

## source

## Location/Qualifiers

```
/clone="cosmid 3A10"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
75308..116845
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
/clone="cosmid 3A10"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
75308..116845
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
```

```
/map="9p13"
/clone="cosmid 3A9"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
/note="BAC 109D8 was subcloned into cosmid8"
114565. .154917
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
/clone="cosmid 3B9"
/clone_1lb="BAC 109D8, Mel Simon's 1lb"
/note="BAC 109D8 was subcloned into cosmid8."
153690. .190172
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
/clone="cosmid X13A"
/clone_1lb="Kai Wang's ATCC1475"
171644. .216293
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9p13"
/clone="cosmid V4G7"
/clone_1lb="Kai Wang's ATCC1475"
27. .148
/rpt_family="AluSx"
151. .288
/rpt_family="L1MA9"
/complement(527. .830)
/rpt_family="AluY"
1152. .2207
/rpt_family="LINE2"
2897. .3496
/rpt_family="L143_Send"
3084. .3495
/rpt_family="L191_Send"
3789. .4395
/rpt_family="L1"
/complement(441. .4460)
/rpt_family="Alu_rich"
4504. .4558
/rpt_family="(CA)n"
/complement(4834. .5238)
/rpt_family="L1ME2"
/complement(5314. .7123)
/note="similar to 10kb tryptsinogen homology unit"
/complement(6042. .6124)
/rpt_family="MIR"
/note="similar to 10 kb tryptsinogen homology unit,
including an orphan exon of tryptsinogen"
7779. .17755
/note="duplication of region from T cell receptor beta
locus on chromosome 7"
/complement(7940. .7965)
/rpt_family="AluYa5"
/complement(8312. .8422)
/rpt_family="(GAAA)n"
9675. .9746
/rpt_family="MER47B"
/complement(9712. .9797)
/rpt_family="MER47A"
9799. .9845
/rpt_family="MER47A"
/complement(10928. .11218)
/rpt_family="AluSg"
11390. .12092
/gene="TCRBV2S20"
```

```
/note="Proposed new name: TCRBV20S2. This pseudogene is
called 'orphan' because it cannot recombine, being on a
different chromosome."
/pseudo
join(11390. .11420,11750. .12053)
/gene="TCRBV2S20"
/standard_name="TCRBV20S2"
/pseudo
12054. .12060
/gene="TCRBV2S20"
/note="RSS heptamer"
12061. .12083
/gene="TCRBV2S20"
/note="RSS spacer"
12084. .12092
/gene="TCRBV2S20"
/note="RSS nonamer"
13007. .13653
/rpt_family="L1ME3"
13007. .13332
/rpt_family="L1ME2"
/complement(13703. .13914)
/rpt_family="MER63A"
14128. .14202
/rpt_family="(TAAA)n"
15161. .15210
/rpt_family="(CA)n"
/complement(15255. .15543)
/rpt_family="AluY"
16413. .16920
/rpt_family="MER34"
/complement(17163. .17311)
/rpt_family="LINE2"
17773. .18894
/note="lipocortin III probable processed pseudogene.
Similar to entries found in Genbank Accession Numbers
D00017 and M62898 and numerous ESTs. Location 17814-18833
gives an in-frame translation."
18891. .19075
/rpt_family="AluSx"
19082. .19370
/rpt_family="AluSx"
19645. .19950
/rpt_family="AluY"

repeat_region
/rpt_family="AluSx"
repeat_region
/rpt_family="AluSx"
repeat_region
/rpt_family="AluY"
misc_feature
/note="lipocortin III probable processed pseudogene.
Similar to entries found in Genbank Accession Numbers
D00017 and M62898 and numerous ESTs. Location 17814-18833
gives an in-frame translation."

Query Match 91.1%; Score 16.4; DB 8; Length 216293;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTCGTGATAC 18
Db 134390 CTCCTCTTCGTGATAC 134373

RESULT 57
AL844559/c 216608 bp DNA linear VRT 28-JAN-2003
LOCUS AL844559
DEFINITION Zebrafish DNA sequence from clone DKX1-14K1, complete sequence.
ACCESSION AL844559
VERSION AL844559.14 GI:28144648
KEYWORDS HTG.
SOURCE Dario rerio (zebrafish)
ORGANISM Dario rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 216608)
REFERENCE
Babbage, A.
Direct Submission
Submitted (28-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 30, 2003 this sequence version replaced gi:27801645.

COMMENT
```



----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [zfish-help@sanger.ac.uk](mailto:zfish-help@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

EMBL, SW: SWISSPROT, Tr: TREMBL, Wp: WORMPEP, Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Chirong Bao and Sean Eddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see [http://www/projects/D\\_retro/fishmask.shtml](http://www/projects/D_retro/fishmask.shtml) DKEY-14K1 is from a Zebrafish BAC library

FEATURES  
source  
Location/Qualifiers

1. 216608  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEY-14K1"  
/clone\_id="DanioKey"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 5; Length 216608;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 1; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGATAC 18  
|||||  
Db 104000 CCCCCCTCTCTGATAC 103983

## RESULT 58

AC113777

LOCUS Rattus norvegicus clone CH230-92115, WORKING DRAFT SEQUENCE, 3  
DEFINITION  
AC113777  
unordered pieces.

AC113777

AC113777.5 GI:25072715

HTG: HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.

Rattus norvegicus (Norway rat)

## SOURCE

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.  
1 (bases 1 to 222869)

REFERENCE

AUTHORS

Murny, D., Marie, Metzger, M., Lee, J., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

## TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

COMMENT

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Bgan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flieger, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Girdly, M., Guerra, W., Guvavara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hu, Y., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensen, L., Loulsegue, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahndarte, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, B., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nockle, M., O., Okunolu, G., Olarinpoju, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Plaz, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., S.J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shastan, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smales, D., Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Treloar, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, J., Wei, X., Wiley, F., Williams, G., Willson, R., Wleczky, R., Woden, H., Wortley, K., Wright, D., Wright, J., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Unpublished

2 (bases 1 to 222869)

Submitted (05-MAR-2002)

Submitted (05-MAR-2002)

Submitted (05-MAR-2002)

Submitted (05-MAR-2002)

Submitted (05-MAR-2002)

Submitted (05-MAR-2002)

Submitted (05-MAR-2002)

Submitted (05-MAR-2002)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 222869)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 19, 2002 this sequence version replaced gi:23269985.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GSGP

```

Center clone name: CH230-92115
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 217955 bases at least Q40
Consensus quality: 218938 bases at least Q30
Consensus quality: 219754 bases at least Q20
Estimated insert size: 223449; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 113340: contig of 113340 bp in length
* 113341 113440: gap of unknown length
* 113441 203688: contig of 90248 bp in length
* 203689 203788: gap of unknown length
* 203789 222869: contig of 19081 bp in length.
FEATURES
source
1..222869
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-92115"
113341..113440
/estimated_length=unknown
113441..115233
/notes="wgs contig"
203689..203788
/estimated_length=unknown
ORIGIN
Query Match 91.1%; Score 16.4; DB 14; Length 222869;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCCCCCTTCGATAC 18
11252 CCCCCCTTCGATAC 11269
RESULT 59
AC103534 225149 bp DNA linear HTG 13-MAY-2003
LOCUS AC103534/c
DEFINITION Rattus norvegicus clone CH230-54A20, *** SEQUENCING IN PROGRESS
ACCESSION AC103534
VERSION AC103534.6 GI:30578685
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
1 (bases 1 to 225149)
Muzny,D,Marle,M, Metzker,M, Lee, A, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Balwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blych, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Devila, M., L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

```

```

Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorjis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Hawlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladik, S., L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Hui, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpanty, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Louised, H., Lozada, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., McNeill, T.Z., Meenen, B.,
Mawliny, S., McLeod, M.P., McNeill, T.Z., Meenen, B.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackemele, O., Okonou, G., Olarunpago, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,
Plapper, F., Polindexter, A., Popovic, D., Primas, E., Pu, L.,
Puzo, M., Qutroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smaj, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soia, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Tjoej, Z., Uman, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleciyk, R., Wooden, R., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 225149)
Worley, K.C.
Direct Submission
Submitted (25-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 225149)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23255269.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GJMW
Center clone name: CH230-54A20
----- Summary Statistics

```



Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information -----  
 Center project name: GJAS  
 Center clone name: CH230-126B18  
 ----- Summary Statistics -----  
 Assembly program: Atlas 3.0  
 Consensus quality: 211082 bases at least Q40  
 Consensus quality: 214888 bases at least Q30  
 Consensus quality: 217080 bases at least Q20  
 Estimated insert size: 223267; sum-of-coverage estimation  
 Quality coverage: 6x in Q20 bases; sum-of-coverage estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 1 226248: contig of 226248 bp in length.  
 Location/Qualifiers  
 1. 226248  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-126B18"  
 55070..57263  
 /note="wgs contig"  
 202670..205731  
 /note="wgs\_contig"

ORIGIN  
 Query Match 91.1%; Score 16.4; DB 14; Length 226248;  
 Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CCCCCCTTTCGATAC 18  
 172439 CCCCCCTTTCGATCC 172422

RESULT 61  
 AC150939/c 237632 bp DNA linear HTG 01-JUL-2005  
 LOCUS Bos taurus clone CH240-512E18, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 DEFINITION unorderd pieces.  
 ACCESSION AC150939  
 VERSION AC150939.2 GI:68265311  
 KEYWORDS HTG; HTGS\_PHSALT; HTGS\_DNAFT; HTGS\_ENRICHED.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
 Pecora; Bovidae; Bovinae; Bos.  
 1 (bases 1 to 237632)  
 Muzny,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,  
 Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,  
 Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,B, Baden,H,  
 Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,  
 Bismio,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,  
 Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,  
 Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A,  
 Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,  
 Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,  
 Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,  
 Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Diya,K,

Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,  
 Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,  
 Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,  
 Frazer,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,  
 Gebregeorgis,B, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,M,  
 Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K,  
 Harey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,  
 Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,  
 Hollins,B, Howells,S, Hulky,S, Hume,J, Idubird,D, Jackson,A,  
 Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolyet,A,  
 Karachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,  
 Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,  
 Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,  
 Lorenshew,L, Loulaeged,H, Lozard,R,J, Lu,X, Ma,J,  
 Maheshwari,M, Mahindarne,M, Mahmood,M, Malloy,K, Mangum,A,  
 Mangum,B, Mapua,P, McNeill,T,Z, Meenen,E,  
 Mawhinney,S, McLeod,M,P, Minja,E, Montemayor,J, Moore,S,  
 Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,  
 Morgan,M, Morris,K, Morris,S, Mundaas,M, Murphy,M, Nair,L,  
 Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,  
 Nwackemele,O, Okwomu,G, Olarunmagbon,A, Pal,S, Parke,K,  
 Pasternak,S, Paul,H, Perez,A, Perez,L, Pfankoch,C,  
 Plopper,F, Polidexter,A, Popovic,D, Primus,E, Pu,L, L,  
 Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,  
 Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,  
 Rivas,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,  
 Sanders,M, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H,  
 Shetty,J, Shvartsbeyn,A, Sison,I, Sitter,C,D, Smajic,D,  
 Sneed,A, Sodergren,E, Song,X-Z, Sorelle,R, Soza,J,  
 Steidle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C,  
 Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umanu,K,  
 Valas,R, Vera,V, Villanasa,D, Waldron,L, Walker,B, Wang,J,  
 Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F,  
 Williams,G, Willson,R, Wleczek,R, Woodem,H, Worley,K,  
 Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V,  
 Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von  
 Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O,  
 Weinstein,G, and Gibbs,R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 237632)  
 Worley,K.C.  
 Direct Submission  
 Submitted (17-AUG-2004) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 237632)  
 Cow Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (01-JUL-2005) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jun 28, 2005 this sequence version replaced gi:51315443.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/ret/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a contig-scaffold). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.  
 ----- Genome Center -----  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information -----  
 Center project name: FBVY  
 Center clone name: CH240-512E18



\* 24700 24799: gap of 100 bp  
\* 24800 34451: contig of 9652 bp in length  
\* 34452 34551: gap of 100 bp  
\* 34552 55796: contig of 21245 bp in length  
\* 55797 55896: gap of 100 bp  
\* 55897 60732: contig of 4835 bp in length  
\* 60732 60832: gap of 100 bp  
\* 60832 72953: contig of 12122 bp in length  
\* 72953 73054: gap of 100 bp  
\* 73054 88799: contig of 15646 bp in length  
\* 88799 88800: gap of 100 bp  
\* 88800 104032: contig of 15223 bp in length  
\* 104032 104132: gap of 100 bp  
\* 104132 120401: contig of 16269 bp in length  
\* 120401 120501: gap of 100 bp  
\* 120501 127691: contig of 7190 bp in length  
\* 127691 12792: gap of 100 bp  
\* 12792 141228: contig of 13437 bp in length  
\* 141228 141329: gap of 100 bp  
\* 141329 148550: contig of 7222 bp in length  
\* 148550 148551: gap of 100 bp  
\* 148551 159671: contig of 11021 bp in length  
\* 159671 159772: gap of 100 bp  
\* 159772 164274: contig of 4503 bp in length  
\* 164274 164375: gap of 100 bp  
\* 164375 170178: contig of 5804 bp in length  
\* 170178 170279: gap of 100 bp  
\* 170279 173659: contig of 3381 bp in length  
\* 173659 173759: gap of 100 bp  
\* 173759 182744: contig of 8984 bp in length  
\* 182744 182844: gap of 100 bp  
\* 182844 196804: contig of 13961 bp in length  
\* 196804 196905: gap of 100 bp  
\* 196905 204024: contig of 7120 bp in length  
\* 204024 204125: gap of 100 bp  
\* 204125 207590: contig of 3466 bp in length  
\* 207590 207691: gap of 100 bp  
\* 207691 222735: contig of 15045 bp in length  
\* 222735 222836: gap of 100 bp  
\* 222836 228372: contig of 5431 bp in length  
\* 228372 228373: gap of 100 bp  
\* 228373 235720: contig of 7348 bp in length  
\* 235720 235821: gap of 100 bp  
\* 235821 238937: contig of 3117 bp in length.

FEATURES

source 1. .238937  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone\_id="RP23-73P8"  
/clone\_lib="RPC1-23"  
1. .7955  
/note="assembly fragment:00289  
fragment\_chain:1  
clone\_end:SP6  
vector\_side:left"  
misc\_feature 8056. .21893  
/note="assembly fragment:01683  
fragment\_chain:1"  
misc\_feature 21994. .34639  
/note="assembly fragment:00568  
fragment\_chain:1"  
misc\_feature 24800. .34451  
/note="assembly fragment:01165  
fragment\_chain:2"  
misc\_feature 34552. .55796  
/note="assembly fragment:02097  
fragment\_chain:2"  
misc\_feature 55897. .60731  
/note="assembly fragment:01095  
fragment\_chain:2"  
misc\_feature 60832. .72953

/note="assembly fragment:01567  
fragment\_chain:2"  
misc\_feature 73054. .88699  
/note="assembly fragment:02092  
fragment\_chain:2"  
misc\_feature 88800. .104032  
/note="assembly fragment:00966  
fragment\_chain:2"  
misc\_feature 104132. .120401  
/note="assembly fragment:02508  
fragment\_chain:3"  
misc\_feature 120502. .127691  
/note="assembly fragment:02483  
fragment\_chain:3"  
misc\_feature 127792. .141228  
/note="assembly fragment:01920  
fragment\_chain:3"  
misc\_feature 141329. .148550  
/note="assembly fragment:01291  
fragment\_chain:3"  
misc\_feature 148551. .159671  
/note="assembly fragment:00807  
fragment\_chain:4"  
misc\_feature 159772. .164274  
/note="assembly fragment:01905  
fragment\_chain:4"  
misc\_feature 164375. .170178  
/note="assembly fragment:00384  
fragment\_chain:5"  
misc\_feature 170279. .173659  
/note="assembly fragment:01356  
fragment\_chain:5"  
misc\_feature 173760. .182743  
/note="assembly fragment:00827  
fragment\_chain:5"  
misc\_feature 182844. .196804  
/note="assembly fragment:01535  
fragment\_chain:5"  
misc\_feature 196905. .204024  
/note="assembly fragment:01678  
fragment\_chain:5"  
misc\_feature 204125. .207590  
/note="assembly fragment:01940  
fragment\_chain:5"  
misc\_feature 207691. .222735  
/note="assembly fragment:00739  
fragment\_chain:5"  
misc\_feature 222836. .228272  
/note="assembly fragment:01503  
fragment\_chain:5"  
misc\_feature 228373. .235720  
/note="assembly fragment:00794  
fragment\_chain:5"  
misc\_feature 235821. .238937  
/note="assembly fragment:01705  
fragment\_chain:5  
clone\_end:17  
vector\_side:right"

ORIGIN

Query Match 91.1% Score 16.4; DB 14; Length 238937;  
Best Local Similarity 94.4%; Pred No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CCCCCCTTCGTGATAC 18  
DB 50748 CCCCCCTTCGTGATAC 50731

RESULT 63  
AC117025 AC117025 247670 bp DNA linear HTG 18-SEP-2002  
LOCUS  
DEFINITION Rattus norvegicus clone CH230-223612, \*\*\* SEQUENCING IN PROGRESS

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

\*\*\*, 4 unordered pieces.  
AC117025  
HTG: HTGS\_PHASE1: HTGS\_DRAFT: HTGS\_ENRICHED.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 247670)

Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguino, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewick, K., Blair, J., Blankenship, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falle, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, L., Louisa, L., Louisa, H., Lorado, R., Lu, X., Ma, J., Maheshwari, M., Mahalingam, M., Mahmood, M., Malloy, K., Mangun, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martineau, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Mlosovlevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Narkervic, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwoko, L., O., Okwuonu, G., Olariu, S., Pal, S., Parke, K., Paeternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poldinger, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Slason, I., Sitter, C.D., Smaja, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taylor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL

2 (bases 1 to 247670)  
Unpublished  
Worley, K.C.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL

3 (bases 1 to 247670)  
Submitted (06-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
Direct Submission  
Rat Genome Sequencing Consortium.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Submitted (18-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Sep 18, 2002 this sequence version replaced gi:21745859. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

## ----- Genome Center

Center: Baylor College of Medicine  
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

## ----- Project Information

Center project name: GIN  
Center clone name: CH230-223G12

## ----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 209112 bases at least Q40  
Consensus quality: 212400 bases at least Q30  
Consensus quality: 214539 bases at least Q20  
Estimated insert size: 258937; sum-of-contigs estimation  
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html))

NOTE: This sequence may represent more than one clone.

NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 204361: contig of 204361 bp in length  
204362 204461: gap of unknown length  
204462 204817: contig of 39356 bp in length  
204818 243918: gap of unknown length  
243918 245571: contig of 1654 bp in length  
245571 245671: gap of unknown length  
245671 247670: contig of 1999 bp in length.

## FEATURES

## Location/Qualifiers

1..247670  
/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-223G12"

204362..204461  
/estimated\_length=unknown

misc\_feature  
/note="wgs contig"

misc\_feature  
/note="wgs contig"

misc\_feature  
/note="wgs contig"

gap  
/estimated\_length=unknown

gap  
/estimated\_length=unknown

gap  
/estimated\_length=unknown

gap  
/estimated\_length=unknown

## ORIGIN

Query Match 91.1%; Score 16.4; DB 14; Length 247670;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTGATAC 18  
AC109565/c CCCCCCTCTGATAC 219122

## RESULT 64

AC109565 251144 bp DNA linear HTG 21-SEP-2002  
LOCUS Rattus norvegicus clone CH230-81G19, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION

ACCESSION AC109565  
 VERSION AC109565.4  
 KEYWORDS HTG: HTGS PHASE2; HTGS\_DRAFT; HTGS\_ENRICHED.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 251144)  
 Muzny, D., Marle, J., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrechts, S., Amin, A., Anguiano, D., Anyadebe, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F., Bismail, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cre, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, M., Gunaratne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulys, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kover, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorusubewa, L., Louised, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, R., Martinez, E., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenan, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleme, O., Okunolu, G., Olarnungsoon, A., Pal, S., Parks, K., Pasceriak, S., Paul, H., Perez, A., Perez, L., Frankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Snead, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usman, K., Valae, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhou, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 251144)  
 Worley, K. C.  
 DIRECT SUBMISSION  
 Submitted (05-PEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 JOURNAL 3 (bases 1 to 251144)  
 Rat Genome Sequencing Consortium.  
 DIRECT SUBMISSION  
 Submitted (21-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Sep 21, 2002 this sequence version replaced gi:21738285. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center -----  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information -----  
 Center project name: G00H  
 Center clone name: CH230-81G19  
 ----- Summary Statistics -----  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 231037 bases at least Q40  
 Consensus quality: 23604 bases at least Q30  
 Consensus quality: 235088 bases at least Q20  
 Estimated insert size: 250342; sum-of-contigs estimation  
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

-----  
 \* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
 \* NOTE: This sequence may represent more than one clone.  
 \* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of 'N'. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 1 251144: contig of 251144 bp in length.  
 Location/Qualifiers  
 source 1..251144  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-81G19"  
 189266..189505  
 misc\_feature  
 /note="clone\_boundary  
 clone\_end:Sp6  
 site:BCORI  
 end\_sequence:BH350518"  
 complement(246448..246636)  
 /note="clone\_boundary  
 clone\_end:T7  
 site:BCORI  
 end\_sequence:BH350517"  
 248279..251144  
 /note="wgs\_end\_extension  
 clone\_end:T7"  
 ORIGIN  
 Query Match 91.1%; Score 16.4; DB 14; Length 251144;  
 Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 CCCCCCTCTTCGATAC 18  
 |||||  
 Db 105542 CCCCCCTCTTCGACAC 105525  
 RESULT 65  
 AC152199/c AC152199 257087 bp DNA linear HTG 01-JUL-2005  
 LOCUS Bos taurus clone CH240-4C2, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 27  
 DEFINITION unordered pieces.  
 ACCESSION AC152199



VERSION AC152199.4 GI.68227172  
 KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
 1 (bases 1 to 257087)  
 Muzny, D., Maric, M., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alibekov, S., Amin, A., Anguiano, D., Anyadebe, C., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, C., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy, Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falle, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebreyes, E., Geier, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, K., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, K., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Maheshwari, M., Mahindaratne, M., Mahmood, M., Martinez, E., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mangum, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Miosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L., Nankervis, C., Neal, D., Newton, G., Olariunasegon, A., Pal, S., Naekeleleh, O., Okunolu, G., Olariunasegon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Plankoch, C., Plopper, F., Poldinger, A., Popovic, D., Primus, E., Pu, L., L., Pizarro, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sibson, I., Sitter, C.D., Smit, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Sytek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleciyk, R., Wood, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE Direct Submission  
 JOURNAL 2 (bases 1 to 257087)  
 AUTHOR WORKLEY, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-OCT-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 257087)  
 REFERENCE Cow Genome Sequencing Consortium.  
 TITLE Submitted (01-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 JOURNAL On Jun 26, 2005 this sequence version replaced gi:58038068.  
 COMMENT The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: FAEQ  
 Center clone name: CH240-4C2  
 ----- Summary Statistics  
 Assembly program: Atlas 3.0!  
 Consensus quality: 248608 bases at least Q40  
 Consensus quality: 250933 bases at least Q30  
 Consensus quality: 252672 bases at least Q20  
 Estimated insert size: 253039; sum-of-contigs estimation  
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

\*\*\*\*\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
 \*\*\*\*\* NOTE: This sequence may represent more than one clone.  
 \*\*\*\*\* NOTE: This is a 'working draft' sequence. It currently consists of 27 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	10132: contig of 10132 bp in length
10133	10182: gap of 50 bp
10183	24465: contig of 14283 bp in length
24466	24515: gap of 50 bp
24516	26754: contig of 2239 bp in length
26755	26804: gap of 50 bp
26805	45962: contig of 19158 bp in length
45963	46012: gap of 50 bp
46013	48982: contig of 2970 bp in length
48983	49032: gap of 50 bp
49033	78973: contig of 29941 bp in length
78974	79023: gap of 50 bp
79024	92341: contig of 13318 bp in length
92342	92391: gap of 50 bp
92392	100051: contig of 7660 bp in length
100052	100101: gap of 50 bp
100102	102699: contig of 2598 bp in length
102700	102749: gap of 50 bp
102750	108516: contig of 5767 bp in length
108517	108566: gap of 50 bp
108567	116746: contig of 8180 bp in length
116747	116796: gap of 50 bp
116797	131322: contig of 14526 bp in length
131323	131372: gap of 50 bp
131373	146337: contig of 14965 bp in length
146338	146387: gap of 50 bp
146388	205665: contig of 59278 bp in length
205666	205715: gap of 50 bp
205716	214313: contig of 8598 bp in length
214314	214363: gap of 50 bp
214364	226352: contig of 11989 bp in length
226353	226402: gap of 50 bp
226403	230250: contig of 3848 bp in length
230251	230349: gap of 99 bp
230350	242925: contig of 12576 bp in length

```

* 242926 243025: gap of unknown length
* 243026 244026: contig of 1001 bp in length
* 244027 244126: gap of unknown length
* 244127 245141: contig of 1015 bp in length
* 245142 245241: gap of unknown length
* 245242 246279: contig of 1038 bp in length
* 246280 246379: gap of unknown length
* 246380 247822: contig of 1443 bp in length
* 247823 247922: gap of unknown length
* 247923 249503: contig of 1581 bp in length
* 249504 249603: gap of unknown length
* 249604 250731: contig of 1128 bp in length
* 250732 250831: gap of unknown length
* 250832 252938: contig of 2107 bp in length
* 252939 253038: gap of unknown length
* 253039 254569: contig of 1531 bp in length
* 254570 257087: gap of unknown length
* 257088 257087: contig of 2418 bp in length.

```

```

FEATURES
source
1..257087
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-4C2"
10133..10182
/estimated_length=50
24466..24515
/estimated_length=50
26755..26804
/estimated_length=50
45963..46012
/estimated_length=50
48983..49032
/estimated_length=50
78974..79023
/estimated_length=50
92342..92391
/estimated_length=50
100052..100101
/estimated_length=50
102700..102749
/estimated_length=50
108517..108566
/estimated_length=50
116747..116796
/estimated_length=50
131323..131372
/estimated_length=50
146338..146387
/estimated_length=50
gap

```

```

Query Match 91.1%; Score 16.4; DB 14; Length 257087;
Best Local Similarity 94.4%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCCCCCTTTTGGATAC 18
DB 241844 CCCCCCTTACTGATAC 241827

```

```

RESULT 66
AC133113/c 258549 bp DNA linear HNG 23-NOV-2002
LOCUS Rattus norvegicus clone CH230-1017, *** SEQUENCING IN PROGRESS ***,
2 unordered pieces.
AC133113
AC133113.3 GI:25188410
ACCESSION HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

```

# REFERENCE

1 (bases 1 to 258549)

Muzny D.M., Adams C., Adio-Oduola B., Ali-ouman F.R., Allen C., Alperbrooks S.L., Amarantunga H.C., Are J.R., Ayele M., Banks T., Barbarella T., Benton J., Binage K., Blankenburg K., Bonin D., Bouck J., Bowe S., Brileva M., Brown E., Brown M., Bryant N.F., Buhay C., Burch S., Burkett C., Burrell K.L., Byrd N.C., Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R., Chen Z., Chowdhury I., Christopoulos C., Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R., Davila M.L., Davis C., Davy-Carroll L., Dederich D.A., Delaney K.R., Delgado O., Dem A.L., Ding Y., Dim H.H., Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J., Earhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M., Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R., Gorell J., H., Guevara W., Gunaratne P., Hale S., Hamilton K., Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J., Hernandez O., Hodgson A., Hogues M., Holloway C., Hollins B., Homel F., Howard S., Huber J., Hulyk S., Hume J., Jackson L.E., Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S., Karlsson E., Kelly S., Khan U., King L., Korvah J., Kovar C., Kratovic V., Kuresh A., Landry N., Leal B., Lewis L.C., Lewis L., Li J., Li Z., Lichtarge O., Lieu C., Liu J., Liu W., Louisaged H., Lozado R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J., Maheshwari M., Mapua P., Martin R., Martindale A., Martinez E., Massey E., Mawhney E., McLeod M.P., Medora M., Mei G., Metzger M., Miner G., Miner Z., Mitchell T., Mohabbat K., Morgan M., Morris S., Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N., Nguyen N., Nickerson E., Nwokenkwo S., Ogun M., Okunolu G., Oragunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L., Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y., Rivers M., Rojas A., Rojibokan I., Rolfe M., Ruiz S., Savary G., Scherer S., Scott G., Shen H., Shooshitari N., Sisson I., Sodergren E., Sonaite T., Sparks A., Stanley H., Stone H., Sutton A., Svatek A., Tabor P., Tamerisa A., Tamerisa K., Tang H., Tansey J., Taylor C., Taylor T., Teliod B., Thomas N., Thomas S., Uemani K., Vaequez L., Vera V., Villalon D., Vinsom R., Wang Q., Wang S., Ward-Moore S., Warren R., Washington C., Watlington S., Williams G., Williamson A., Wleczek R., Wooden S., Worley K., Wu C., Wu Y., Wu Y.P., Zhou J., Zorilla S., Nelson D., Weinstein G., and Gibbs R.

# TITLE

Unpublished  
2 (bases 1 to 258549)  
Worley K.C.  
Direct Submission  
Submitted (07-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 258549)  
Worley K.C.  
Direct Submission  
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

# REFERENCE

Unpublished  
2 (bases 1 to 258549)  
Worley K.C.  
Direct Submission  
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

# COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu

```

----- Project Information
Center project name: TUMQ
Center clone name: CH230-1017
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 206883 bases at least Q40
Consensus quality: 213001 bases at least Q30
Consensus quality: 217247 bases at least Q20
Estimated insert size: 211792; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 257128: contig of 257128 bp in length
* 257129 258549: gap of unknown length
* 257229 258549: contig of 1321 bp in length.
Location/Qualifiers
1. 258549
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-1017"
201389..202531
/notes="wgs contig"
220604..221612
/notes="wgs contig"
257129..257228
/estimated_length=unknown

FEATURES
source
1. 258549
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-1017"
201389..202531
/notes="wgs contig"
220604..221612
/notes="wgs contig"
257129..257228
/estimated_length=unknown

ORIGIN
Query Match 91.1%; Score 16.4; DB 14; Length 258549;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCCCCCTTCTGTGATAC 18
Db 93668 CCCCCCTTCTGTGACAC 93651

RESULT 67
AC118292
LOCUS
DEFINITION
Rattus norvegicus clone CH230-212F24, *** SEQUENCING IN PROGRESS
*** 7 unordered pieces.
AC118292
AC118292.4 GI:22856547
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 258819)
Muzny,D,Marie., Metzger,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alabrooks,S., Amin,A., Angiano,D.,
Auyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalto,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryan,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., P.Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,J., De Anda,C., Dedrich,D.,
Deigado,O., Denson,S., Derramo,C., Ding,Y., Dinh,H., Diya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTR8
Center clone name: CH230-212F24
----- Summary Statistics
Egan,A., Baccotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabrieli,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Grebegeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,B., Hines,S., Hladun,S., Hodgson,A., Hogue,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,J., Kovar,C.,
Kowitz,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenznew,L., Louised,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmud,M., Mallory,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenan,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,C., Moore,S.,
Morgan,M., Morris,K., Morris,S., Muntadas,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackelameh,O., Okwunonu,G., Olarnungsoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Saverly,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,A., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajz,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorrell,R., Soes,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabors,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valae,R., Vera,V., Villasea,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhou,S., Dunn,D., von
Niederhausen,A., Weiser,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 258819)
Worley,K.C.
Direct Submission
Submitted (15-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 258819)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 14, 2002 this sequence version replaced gi:21745807.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTR8
Center clone name: CH230-212F24
----- Summary Statistics

```

Assembly program: Phrap, version 0.990329  
Consensus quality: 186841 bases at least Q40  
Consensus quality: 191174 bases at least Q30  
Consensus quality: 194141 bases at least Q20  
Estimated insert size: 207863; sum-of-contigs estimation  
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Gendbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Gendbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 247432: contig of 247432 bp in length  
\* 247433 247532: gap of unknown length  
\* 247533 248891: contig of 1359 bp in length  
\* 248891 248991: gap of unknown length  
\* 248991 250023: contig of 1032 bp in length  
\* 250023 250123: gap of unknown length  
\* 250123 251753: contig of 1630 bp in length  
\* 251753 251853: gap of unknown length  
\* 251853 254088: contig of 2235 bp in length  
\* 254088 254188: gap of unknown length  
\* 254188 255707: contig of 1519 bp in length  
\* 255707 255808: gap of unknown length  
\* 255808 258819: contig of 3012 bp in length.  
\*  
Location/Qualifiers  
1..258819 Rattus norvegicus"  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-212P24"  
98807..100754  
/note="wgs\_contig"  
182556..183608  
/note="wgs\_contig"  
247433..247532  
/estimated\_length=unknown  
248891..248991  
/estimated\_length=unknown  
250023..250123  
/estimated\_length=unknown  
251754..251853  
/estimated\_length=unknown  
254089..254188  
/estimated\_length=unknown  
255708..255807  
/estimated\_length=unknown  
-----  
ORIGIN  
Query Match 91.1%; Score 16.4; DB 14; Length 258819;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CCCCCCTTCTGGATAC 18  
|||  
Db 147119 CCCCCCTTCTGGATAC 147136  
|||  
RESULT 68  
AL929047 261202 bp DNA linear HTG 24-JAN-2003  
LOCUS Mus musculus chromosome 2 clone RP24-250D1, 2 unordered pieces.  
ACCESSION AL929047 GI:27899643  
VERSION AL929047.5  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_CANCELLED.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurionanthi; Muroidae; Muridae; Murinae; Mus.  
1 (bases 1 to 261202)  
Tracey/A.  
Direct Submission  
Submitted (23-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Jan 24, 2003 this sequence version replaced gi:27858301.  
Sequence from the Mouse Genome Sequencing Consortium whole genome  
shotgun may have been used to confirm this sequence. Sequence data  
from the whole genome shotgun alone has only been used where it has  
a phred quality of at least 30.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: humquerry@sanger.ac.uk  
----- Project Information  
Center project name: bn250D1  
----- Summary Statistics  
Assembly program: XGAP4, version 4.5  
Chemistry: dye-terminator, 100% of reads  
Consensus quality: 180099 bases at least Q40  
Consensus quality: 180231 bases at least Q30  
Consensus quality: 180268 bases at least Q20  
Insert size: 261102; sum-of-contigs  
Insert size: 186286; 6.1% error; agarose-fp  
Quality coverage: 5.59x in Q20 bases; sum-of-contigs Quality  
coverage: 7.87x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 231177: contig of 231177 bp in length  
\* 231178 231278: gap of 100 bp  
\* 231278 261202: contig of 29925 bp in length.  
\*  
Location/Qualifiers  
1..261202  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="2"  
/clone="RP24-250D1"  
/clone\_1fb="RPCL-24"  
1..231177  
/note="assembly\_fragment:02118  
fragment\_chain:1  
clone\_end:17  
vector\_side:left"  
231278..261202  
/note="assembly\_fragment:00575  
fragment\_chain:1"  
-----  
ORIGIN  
Query Match 91.1%; Score 16.4; DB 14; Length 261202;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CCCCCCTTCTGGATAC 18  
|||  
Db 79549 CCCCCCTTCTGGATAC 79566  
|||  
RESULT 69  
AC104926 285603 bp DNA linear HTG 03-MAR-2003  
LOCUS Mus musculus clone RP24-573F21, WORKING DRAFT SEQUENCE, 12  
DEFINITION

# ACCESSION VERSION KEYWORDS SOURCE ORGANISM

unordered pieces.  
AC104926  
AC104926.5 GI:28631311  
HTG; HTGS PHASE1; HTGS DRAFT.  
Mus musculus (house mouse)

# REFERENCE AUTHORS TITLE JOURNAL REFERENCES AUTHORS

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 285603)  
Birren, B., Nuebaum, C. and Lander, E.  
Mus musculus, clone RP24-573P21  
Unpublished  
2 (bases 1 to 285603)

Birren, B., Linton, L., Nuebaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouhgalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J., Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, U.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talmas, J., Testaye, S., Thodode, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

# TITLE JOURNAL REFERENCE AUTHORS

Direct Submission  
Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 285603)

Birren, B., Nuebaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, U., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hages, B., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talmas, J., Testaye, S., Theodore, J., Topham, K., Travers, B., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, U., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (03-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 3, 2003 this sequence version replaced gi:25989100.  
All repeats were identified using RepeatMasker:  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WITBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information

Center project name: L20112  
Center file name: 573\_P\_21  
----- Summary Statistics -----  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 281940 bases at least Q40  
Consensus quality: 283518 bases at least Q30  
Consensus quality: 284159 bases at least Q20  
Insert size: 24800; agarose-fp  
Insert size: 284503; sum-of-contigs  
Quality coverage: 10.1 in Q20 bases; agarose-fp  
Quality coverage: 8.8 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 1204: contig of 1204 bp in length  
\* 1205 1304: gap of 100 bp  
\* 1305 4414: contig of 3110 bp in length  
\* 4415 4514: gap of 100 bp  
\* 4515 7932: contig of 3418 bp in length  
\* 7933 8032: gap of 100 bp  
\* 8033 12101: contig of 4069 bp in length  
\* 12102 12201: gap of 100 bp  
\* 12202 16143: contig of 3942 bp in length  
\* 16144 16243: gap of 100 bp  
\* 16244 21154: contig of 4911 bp in length  
\* 21155 21254: gap of 100 bp  
\* 21255 25591: contig of 4337 bp in length  
\* 25592 25692: gap of 100 bp  
\* 25693 40122: contig of 14431 bp in length  
\* 40123 40222: gap of 100 bp  
\* 40223 59108: contig of 18886 bp in length  
\* 59109 59208: gap of 100 bp  
\* 59209 189096: contig of 129888 bp in length  
\* 189097 189196: gap of 100 bp  
\* 189197 274853: contig of 85657 bp in length  
\* 274854 274953: gap of 100 bp  
\* 274954 285603: contig of 10650 bp in length.  
\*  
FEATURES  
source  
1. 285603  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP24-573P21"  
/clone\_lib="RPCI-24 Male Mouse BAC"  
1. 1204  
/note="assembly\_fragment"  
1205. 1304  
/estimated\_length=100  
1305. 4414  
/note="assembly\_fragment"  
4415. 4514  
/estimated\_length=100  
4515. 7932  
/note="assembly\_fragment"  
7933. 8032  
/estimated\_length=100  
8033. 12101  
/note="assembly\_fragment"  
12102. 12201  
/estimated\_length=100  
12202. 16143  
/note="assembly\_fragment"  
16144. 16243  
/estimated\_length=100  
16244. 21154  
misc\_feature  
1204: contig of 1204 bp in length  
misc\_feature  
1205: gap of 100 bp  
misc\_feature  
1305: gap of 100 bp  
misc\_feature  
4415: gap of 100 bp  
misc\_feature  
4515: gap of 100 bp  
misc\_feature  
7933: gap of 100 bp  
misc\_feature  
8033: gap of 100 bp  
misc\_feature  
12102: gap of 100 bp  
misc\_feature  
12202: gap of 100 bp  
misc\_feature  
16144: gap of 100 bp  
misc\_feature  
16244: gap of 100 bp  
misc\_feature  
21155: gap of 100 bp  
misc\_feature  
25592: gap of 100 bp  
misc\_feature  
40123: gap of 100 bp  
misc\_feature  
40223: gap of 100 bp  
misc\_feature  
59109: gap of 100 bp  
misc\_feature  
189097: gap of 100 bp  
misc\_feature  
189197: gap of 100 bp  
misc\_feature  
274854: gap of 100 bp  
misc\_feature  
274954: gap of 100 bp  
misc\_feature  
285603: gap of 100 bp

gap /note="assembly\_fragment"  
21155..21254  
/estimated\_length=100  
misc\_feature 21255..25591  
/note="assembly\_fragment"  
25592..25691  
/estimated\_length=100  
misc\_feature 25692..40122  
/note="assembly\_fragment"  
40123..40222  
/estimated\_length=100  
misc\_feature 40223..59108  
/note="assembly\_fragment"  
59109..59208  
/estimated\_length=100  
misc\_feature 59209..189096  
/note="assembly\_fragment"  
189097..189196  
/estimated\_length=100  
misc\_feature 189197..274853  
/note="assembly\_fragment"  
274854..274953  
/estimated\_length=100  
gap 274954..285603  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right"

ORIGIN

Query Match 91.1%; Score 16.4; DB 14; Length 285603;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18  
Db 238010 CCCCCCTTCTGGATAC 237993

RESULT 70 CR376752 344922 bp DNA 1linear HTG 11-JAN-2005  
LOCUS CR376752/c  
DEFINITION Danio rerio clone CH211-201P11, WORKING DRAFT SEQUENCE, 10  
unordered pieces.  
ACCESSION CR376752  
VERSION CR376752.7 GI:57506444  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVEPIN; HTGS\_DRAFT; HTGS\_FULITOP.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 344922)  
Almeida,J.  
Direct Submission  
Submitted (10-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Jan 11, 2005 this sequence version replaced gi:52213867.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: zfish-help@sanger.ac.uk  
----- Project Information  
Center project name: ZC201P11  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 196056 bases at least Q40  
Consensus quality: 197521 bases at least Q30  
Consensus quality: 198625 bases at least Q20  
Insert size: 344022; sum-of-contigs  
Insert size: 179984; 1.6% error; agarose-fp

Quality coverage: 4.28x in Q20 bases; sum-of-contigs Quality  
coverage: 8.68x in Q20 bases; agarose-fp

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 71996: contig of 71996 bp in length  
\* 71997 72096: gap of 100 bp  
\* 72097 81126: contig of 9030 bp in length  
\* 81127 81226: gap of 100 bp  
\* 81227 83753: contig of 2527 bp in length  
\* 83754 83853: gap of 100 bp  
\* 83854 89809: contig of 5956 bp in length  
\* 89810 89909: gap of 100 bp  
\* 89910 99771: contig of 9662 bp in length  
\* 99772 99871: gap of 100 bp  
\* 99872 178769: contig of 78898 bp in length  
\* 178770 178869: gap of 100 bp  
\* 178870 182663: contig of 3794 bp in length  
\* 182664 182763: gap of 100 bp  
\* 182764 188792: contig of 6029 bp in length  
\* 188793 201134: gap of 100 bp  
\* 201135 201234: contig of 12242 bp in length  
\* 201235 344922: gap of 100 bp  
\* 201235 344922: contig of 143688 bp in length.

FEATURES

source

misc\_feature 1..344922  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="CH211-201P11"  
/clone\_1fb="CHORI-211"  
1..71996  
/note="assembly\_fragment:00832  
clone\_end:SP6  
vector\_side:left"  
72097..81126  
/note="assembly\_fragment:00376  
fragment\_chain:1"  
81227..83753  
/note="assembly\_fragment:00063  
fragment\_chain:1"  
83854..89809  
/note="assembly\_fragment:00974  
fragment\_chain:1"  
89910..99771  
/note="assembly\_fragment:00377  
fragment\_chain:1"  
99872..178769  
/note="assembly\_fragment:01651  
fragment\_chain:1"  
178870..182663  
/note="assembly\_fragment:00026"  
182764..188792  
/note="assembly\_fragment:00040"  
188893..201134  
/note="assembly\_fragment:00100"  
201235..344922  
/note="assembly\_fragment:02748"  
ORIGIN  
Query Match 91.1%; Score 16.4; DB 14; Length 344922;  
Best Local Similarity 94.4%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCCCCCTTCTGGATAC 18  
Db 303337 CCCCCCTTCTGGATAC 303320

Mon May 8 11:07:23 2006

us-10-659-980a-5.rge

Page 49

Search completed: May 5, 2006, 15:04:41  
Job time : 1769 secs

11:07:23 AM  
May 8, 2006  
us-10-659-980a-5.rge

**THIS PAGE LEFT BLANK**



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 6, 2006, 15:49:03 ; Search time 294 Seconds  
(without alignments)  
408.042 Million cell updates/sec

Title: US-10-659-980a-5

Perfect score: 18

Sequence: 1 cccccctctcgtgatac 18

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 70 summaries

Database:

1: N\_Geneseq\_21.\*  
2: geneseq1980s.\*  
3: geneseq1990s.\*  
4: geneseq2000s.\*  
5: geneseq2001bs.\*  
6: geneseq2002as.\*  
7: geneseq2002bs.\*  
8: geneseq2003as.\*  
9: geneseq2003bs.\*  
10: geneseq2003cs.\*  
11: geneseq2003ds.\*  
12: geneseq2004as.\*  
13: geneseq2004bs.\*  
14: geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	6	ABA02420 Type A/Al
2	18	100.0	18	12	ADM32708 Nitrosp
3	18	100.0	1457	6	ABA02416 Type A am
4	18	100.0	1457	6	ABA02417 Type A1 a
5	18	100.0	1457	12	ADM32705 AOB Type
6	18	100.0	1457	12	ADM32704 AOB Type
7	16	88.9	289	11	ADM65290 NRY polym
8	15.4	85.6	448	13	ACF84919 Human SIR
9	15.4	85.6	486	13	ACF83823 Human SIR
10	15.4	85.6	1104	8	ACA27476 Prokaryot
11	15.4	85.6	1866	2	AAO04292 Encodes p
12	15.4	85.6	6388	7	ADY36534 HIRA geno
13	15.4	85.6	6388	7	ADY36534 HIRA geno
14	15.4	85.6	12542	4	ABL12362 Drosophil
15	15.4	85.6	67810	11	ACN45006 Human gen
16	15.4	85.6	79256	13	ABD31177 Murine ca
17	15.4	85.6	110000	10	ADH10017 1
18	15.4	85.6	110000	12	ADQ97266_0
19	15.4	85.6	110000	12	ADQ97266_1

## ALIGNMENTS

20	15	83.3	409	9	ACH28985	ACH28985 Human adu
21	15	83.3	2300	10	ADG31675	Adg31675 Human CDN
22	15	83.3	2405	4	AAO02351	AAO02351 Human RNA
23	15	83.3	2776	10	ADB62993	ADB62993 Human CDN
24	15	83.3	2856	10	ADG31677	Adg31677 Human CDN
25	15	83.3	5208	13	ACN39725	ACN39725 Tumour-as
26	15	83.3	5581	10	ADK65780	ADK65780 Angiogene
27	15	83.3	5647	6	ABO61143	ABO61143 Nucleic p
28	15	83.3	6367	10	ADC30216	Adc30216 Human nov
29	15	83.3	60989	10	ADC85744	Adc85744 Human GPC
30	15	83.3	110000	13	ABD32627_0	ABD32627 Mouse can
31	15	83.3	224931	14	AD209671	Ad209671 Human bre
32	14.8	82.2	413	10	ADK53800	Adk53800 Human pro
33	14.8	82.2	616	13	ADQ79063	Adq79063 Novel can
34	14.8	82.2	682	4	AA533271	AA533271 DNA encod
35	14.8	82.2	682	4	AA533346	AA533346 DNA encod
36	14.8	82.2	682	4	AA185744	AA185744 Human pol
37	14.8	82.2	1095	4	ABL26119	AbL26119 Drosophil
38	14.8	82.2	1157	5	AA592387	AA592387 DNA encod
39	14.8	82.2	1408	3	AAQ42079	AAQ42079 Arabidops
40	14.8	82.2	2297	13	ADT19869	Adt19869 plant CDN
41	14.8	82.2	2433	5	AA579005	AA579005 DNA encod
42	14.8	82.2	3095	4	ABL26118	AbL26118 Novel hum
43	14.8	82.2	3788	12	ADQ63840	AdQ63840 Human imm
44	14.8	82.2	3861	4	AAK73805	Aak73805 Human imm
45	14.8	82.2	3861	4	AAK73521	Aak73521 Human imm
46	14.8	82.2	6212	14	ADY16196	Ady16196 DNA encod
47	14.8	82.2	9830	3	AAV49655	AAV49655 Nucleotid
48	14.8	82.2	28720	2	AAV49655	AAV49655 Human SC3
49	14.8	82.2	30135	13	ACN37225	ACN37225 Human per
50	14.8	82.2	69648	12	ADQ97934	Adq97934 Mouse gen
51	14.8	82.2	82660	11	ACN45192	ACN45192 Mouse gen
52	14.8	82.2	83493	14	ADZ13310	Adz13310 Murine ca
53	14.8	82.2	109147	12	ADQ97718	AdQ97718 Mouse can
54	14.8	82.2	110000	12	ADN46845_18	ADN46845_18 o
55	14.8	82.2	110000	12	ADN47591_02	ADN47591_02 o
56	14.8	82.2	110000	12	ADN47209_02	ADN47209_02 o
57	14.8	82.2	110000	12	ADN46464_18	ADN46464_18 o
58	14.8	82.2	110000	12	ADN47960_02	ADN47960_02 o
59	14.8	82.2	174566	8	ABQ77400	AbQ77400 Human ITG
60	14.8	82.2	174566	12	ADL08118	AdL08118 Human gen
61	14.8	82.2	260803	13	ABD32730	ABD32730 Human can
62	14.8	82.2	260803	13	AA511828	AA511828 Human sur
63	14.4	80.0	30	4	AA511828	AA511828 Human bre
64	14.4	80.0	276	4	AA115054	AA115054 Human bre
65	14.4	80.0	307	4	AA191862	AA191862 Human pol
66	14.4	80.0	370	6	ABO56427	ABO56427 Human col
67	14.4	80.0	378	6	ABN16125	ABN16125 Human ORF
68	14.4	80.0	393	8	ABX64829	ABX64829 Human gen
69	14.4	80.0	406	10	ADP59862	Adp59862 Human con
70	14.4	80.0	419	6	ABL80291	AbL80291 Human ova

RESULT 1  
ABA02420  
ID ABA02420 standard; DNA; 18 BP.

29-AUG-2003 (revised)  
04-MAR-2002 (first entry)

Type A/Al ammonia-oxidising bacterium 16S rRNA gene probe.

Type A: type A1; ammonia-oxidising bacterium; AOB; nitrite;

16S rRNA gene; ribosomal RNA; aquaculture;

waste water treatment; bioremediation; probe; ss.

Nitrosomonadaceae.

PN WO200190312-A1.  
 XX  
 XX 29-NOV-2001.  
 PD  
 XX 17-MAY-2001; 2001WO-US016265.  
 PF  
 XX 19-MAY-2000; 2000US-00573684.  
 PR  
 XX (AQUA-) AQUARIA INC.  
 PA  
 XX Hovanec TA, Burrell PC;  
 PI  
 XX MPI; 2002-075367/10.  
 DR  
 XX  
 XX New bacteria capable of oxidizing ammonia to nitrite, for preventing or  
 PT alleviating the accumulation of ammonia in fresh water aquaria, seawater  
 PT aquaria and waste water.  
 PT  
 XX  
 PS Claim 26; Page 53; 62pp; English.  
 CC  
 XX The invention relates to 4 novel types of ammonia-oxidizing bacteria  
 CC (AOB) found in freshwater aquaria. The bacteria are able to oxidise  
 CC ammonia to nitrite and are members of the ammonia-oxidising bacteria  
 CC family of the beta subdivision of Proteobacteria. The 4 types of bacteria  
 CC can be distinguished on the basis of their 16S rRNA (ribosomal RNA) gene  
 CC sequences (ABA02416-ABA02419), and are classified as AOB type A (e.g.,  
 CC R7clone140), type A1 (e.g., R7clone187), type B (e.g., R3clone5) and type  
 CC C (e.g., R3clone47). The invention also encompasses isolated 16S rRNA  
 CC oligonucleotide probes and primers for the detection of these bacteria,  
 CC and compositions comprising the bacteria. The bacteria of the invention  
 CC are useful in biological filters for reducing ammonia accumulation in  
 CC both freshwater and seawater aquaria. They may also be used in waste  
 CC water treatment and in bioremediation processes to reduce the level of  
 CC pollution caused by ammonia. The present sequence represents a  
 CC specifically claimed probe for the detection of the 16S rRNA gene  
 CC sequences of the type A and A1 ammonia-oxidising bacteria (ABA02416-  
 CC ABA02417). (Updated on 29-AUG-2003 to standardise OS field)  
 CC  
 XX  
 SQ Sequence 18 BP; 2 A; 9 C; 2 G; 5 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 18; DB 6; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CCCCCCTCTTCGTGATAC 18  
 DB 1 CCCCCCTCTTCGTGATAC 18  
 RESULT 2  
 ADM32708  
 ID ADM32708 standard; DNA; 18 BP.  
 XX  
 AC ADM32708;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Nitrosospira-like 16S rDNA probe. S-G-Nsapa-0149-a-A-18.  
 XX  
 DE Nitrosospira-like 16S rDNA probe. S-G-Nsapa-0149-a-A-18.  
 XX  
 KM 16S rDNA; ammonia oxidising bacteria; AOB; ammonia; nitrite;  
 KM aqueous environment; freshwater; seawater; aquarium; sea probe.  
 KM  
 OS Nitrosospira sp.  
 XX  
 XX WO2004026772-A2.  
 PN  
 XX  
 PD 01-APR-2004.  
 XX  
 PF 10-SEP-2003; 2003WO-US028210.  
 XX  
 PR 19-SEP-2002; 2002US-0386217P.  
 XX  
 PR 19-SEP-2002; 2002US-0386218P.

PR 19-SEP-2002; 2002US-0386219P.  
 XX  
 XX (AQUA-) AQUARIA INC.  
 PA  
 XX Hovanec TA;  
 PI  
 XX MPI; 2004-304936/28.  
 DR  
 XX  
 XX New composition comprising an isolated bacterial strain that oxidizes  
 PT ammonia to nitrite, useful for alleviating or preventing the accumulation  
 PT of ammonia in aqueous environment.  
 PT  
 XX  
 PS Example 6; Page 34; 98pp; English.  
 CC  
 XX This sequence represents a probe which was used in the detection of 16S  
 CC rDNA sequence derived from an ammonia oxidising bacteria (AOB). The  
 CC amplified sequence may be used in a composition which comprises an  
 CC isolated bacterial strain that oxidizes ammonia to nitrite. The  
 CC composition may be used for alleviating or preventing the accumulation of  
 CC ammonia in a medium. The ammonia is reduced by at least 30% when compared  
 CC with a level of ammonia that would exist in the absence of the bacterial  
 CC strain. The composition is useful for alleviating or preventing the  
 CC accumulation of ammonia in aqueous environment, e.g. a freshwater or  
 CC seawater aquarium.  
 CC  
 XX  
 SQ Sequence 18 BP; 2 A; 9 C; 2 G; 5 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 18; DB 12; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CCCCCCTCTTCGTGATAC 18  
 DB 1 CCCCCCTCTTCGTGATAC 18  
 RESULT 3  
 ABA02416/C  
 ID ABA02416 standard; DNA; 1457 BP.  
 XX  
 AC ABA02416;  
 XX  
 DT 29-AUG-2003 (revised)  
 DT 04-MAR-2002 (first entry)  
 XX  
 DE Type A ammonia-oxidising bacterium 16S rRNA gene sequence. R7clone140.  
 XX  
 KM Type A; ammonia-oxidising bacterium; AOB; nitrite; 16S rRNA gene;  
 KM ribosomal RNA; aquarium; aquaculture; waste water treatment;  
 KM bioremediation; ds.  
 KM  
 OS Nitrosomonadales.  
 OS  
 XX  
 XX WO200190312-A1.  
 PN  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 17-MAY-2001; 2001WO-US016265.  
 XX  
 PR 19-MAY-2000; 2000US-00573684.  
 XX  
 PA (AQUA-) AQUARIA INC.  
 XX  
 XX Hovanec TA, Burrell PC;  
 PI  
 XX MPI; 2002-075367/10.  
 DR  
 XX  
 XX New bacteria capable of oxidizing ammonia to nitrite, for preventing or  
 PT alleviating the accumulation of ammonia in fresh water aquaria, seawater  
 PT aquaria and waste water.  
 PT  
 XX  
 PS Claim 2; Page 5; 62pp; English.

CC The invention relates to 4 novel types of ammonia-oxidising bacteria  
CC (AOB) found in freshwater aquaria. The bacteria are able to oxidise  
CC ammonia to nitrite and are members of the ammonia-oxidising bacteria  
CC family of the beta subdivision of Proteobacteria. The 4 types of bacteria  
CC can be distinguished on the basis of their 16S rRNA (ribosomal RNA) gene  
CC sequences (ABA02416-ABA02419), and are classified as AOB type A (e.g.,  
CC R7clone140), type A1 (e.g., R7clone187), type B (e.g., R3clone5) and type  
CC C (e.g., R3clone47). The invention also encompasses isolated 16S rRNA  
CC gene sequences of the ammonia-oxidising bacteria of the invention,  
CC oligonucleotide probes and primers for the detection of these bacteria,  
CC and compositions comprising the bacteria. The bacteria of the invention  
CC are useful in biological filters for reducing ammonia accumulation in  
CC both freshwater and seawater aquaria. They may also be used in waste  
CC water treatment and in bioremediation processes to reduce the level of  
CC pollution caused by ammonia. The present sequence represents R7clone140,  
CC a 16S rRNA gene sequence from the type A ammonia-oxidising bacterium of  
CC the invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 1457 BP; 378 A; 315 C; 456 G; 308 T; 0 U; 0 Other;

SO Query Match 100.0%; Score 18; DB 6; Length 1457;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGGATAC 18

DB 118 CCCCCCTCTCTGGATAC 101

RESULT 4  
ABA02417/c  
ID ABA02417 standard; DNA; 1457 BP.

AC ABA02417;

XX 29-AUG-2003 (revised)

DT 04-MAR-2002 (first entry)

XX Type A1 ammonia-oxidising bacterium 16S rRNA gene sequence, R7clone187.

XX Type A1; ammonia-oxidising bacterium; AOB; nitrite; 16S rRNA gene;

KM ribosomal RNA; aquarium; aquaculture; waste water treatment;

KM bioremediation; ds.

XX Nitrosomonadales.

XX WO200190312-A1.

XX 29-NOV-2001.

XX 17-MAY-2001; 2001WO-US016265.

XX 19-MAY-2000; 2000US-00573684.

XX (AQUA-) AQUARIA INC.

XX Hovanec TA, Burrell PC;

XX WPI; 2002-075367/10.

XX New bacteria capable of oxidizing ammonia to nitrite, for preventing or

XX alleviating the accumulation of ammonia in fresh water aquaria, seawater

XX aquaria and waste water.

XX Claim 2; Page 6; 62pp; English.

XX The invention relates to 4 novel types of ammonia-oxidising bacteria

XX (AOB) found in freshwater aquaria. The bacteria are able to oxidise

XX ammonia to nitrite and are members of the ammonia-oxidising bacteria

XX family of the beta subdivision of Proteobacteria. The 4 types of bacteria

XX can be distinguished on the basis of their 16S rRNA (ribosomal RNA) gene

XX sequences (ABA02416-ABA02419), and are classified as AOB type A (e.g.,

XX R7clone140), type A1 (e.g., R7clone187), type B (e.g., R3clone5) and type

CC C (e.g., R3clone47). The invention also encompasses isolated 16S rRNA  
CC gene sequences of the ammonia-oxidising bacteria of the invention,  
CC oligonucleotide probes and primers for the detection of these bacteria,  
CC and compositions comprising the bacteria. The bacteria of the invention  
CC are useful in biological filters for reducing ammonia accumulation in  
CC both freshwater and seawater aquaria. They may also be used in waste  
CC water treatment and in bioremediation processes to reduce the level of  
CC pollution caused by ammonia. The present sequence represents R7clone187,  
CC a 16S rRNA gene sequence from the type A1 ammonia-oxidising bacterium of  
CC the invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 1457 BP; 379 A; 317 C; 455 G; 306 T; 0 U; 0 Other;

SO Query Match 100.0%; Score 18; DB 6; Length 1457;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGGATAC 18

DB 118 CCCCCCTCTCTGGATAC 101

RESULT 5  
ADM32705/c  
ID ADM32705 standard; DNA; 1457 BP.

AC ADM32705;

XX 17-JUN-2004 (first entry)

XX AOB Type A1 R7clone187 16S rDNA.

XX 16S rDNA; ammonia oxidising bacteria; AOB; ammonia; nitrite;

KM aqueous environment; freshwater; seawater; aquarium; ss.

XX Nitrosomonas sp.

XX WO2004026772-A2.

XX 01-APR-2004.

XX 10-SEP-2003; 2003WO-US028210.

XX 19-SEP-2002; 2002US-0386217P.

XX 19-SEP-2002; 2002US-0386218P.

XX 19-SEP-2002; 2002US-0386219P.

XX (AQUA-) AQUARIA INC.

XX Hovanec TA;

XX WPI; 2004-304936/28.

XX New composition comprising an isolated bacterial strain that oxidizes

XX ammonia to nitrite, useful for alleviating or preventing the accumulation

XX of ammonia in aqueous environment.

XX Disclosure; Page 10; 98pp; English.

XX This sequence represents a 16S rDNA sequence derived from an ammonia

XX oxidising bacteria (AOB). This sequence may be used in a composition

XX which comprises an isolated bacterial strain that oxidizes ammonia to

XX nitrite. The composition may be used for alleviating or preventing the

XX accumulation of ammonia in a medium. The ammonia is reduced by at least

XX 30% when compared with a level of ammonia that would exist in the absence

XX of the bacterial strain. The composition is useful for alleviating or

XX preventing the accumulation of ammonia in aqueous environment, e.g., a

XX freshwater or seawater aquarium.

SO Sequence 1457 BP; 379 A; 317 C; 455 G; 306 T; 0 U; 0 Other;

Query Match 100.0%; Score 18; DB 12; Length 1457;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCTTTGTGATAC 18  
118 CCCCCCTTTGTGATAC 101

RESULT 6  
ADM32704/c  
ID ADM32704 standard; DNA; 1457 BP.

AC ADM32704;

DT 17-JUN-2004 (first entry)

DE AOB Type A R7c1one140 16S rDNA.

XX 16S rDNA: ammonia oxidizing bacteria; AOB; ammonia; nitrite;

KM aqueous environment; freshwater; seawater; aquarium; ss.

XX Nitrosomonas sp.

PN WO2004026772-A2.

XX 01-APR-2004.

PF 10-SEP-2003; 2003WO-US028210.

PR 19-SEP-2002; 2002US-0386217P.

PR 19-SEP-2002; 2002US-0386218P.

XX 19-SEP-2002; 2002US-0386219P.

PA (AQUA-) AQUARIA INC.

XX Hovanec TA;

DR WPI; 2004-304936/28.

PT New composition comprising an isolated bacterial strain that oxidizes

PT ammonia to nitrite, useful for alleviating or preventing the accumulation

PT of ammonia in aqueous environment.

XX Disclosure; Page 8-9; 98pp; English.

CC This sequence represents a 16S rDNA sequence derived from an ammonia

CC oxidizing bacteria (AOB). This sequence may be used in a composition

CC which comprises an isolated bacterial strain that oxidizes ammonia to

CC nitrite. The composition may be used for alleviating or preventing the

CC accumulation of ammonia in a medium. The ammonia is reduced by at least

CC 30% when compared with a level of ammonia that would exist in the absence

CC of the bacterial strain. The composition is useful for alleviating or

CC preventing the accumulation of ammonia in aqueous environment, e.g. a

CC freshwater or seawater aquarium.

XX Sequence 1457 BP; 378 A; 315 C; 456 G; 308 T; 0 U; 0 Other;

QY Query Match 100.0%; Score 18; DB 12; Length 1457;

AC Best Local Similarity 100.0%; Pred. No. 60;

XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 118 CCCCCCTTTGTGATAC 101

RESULT 7  
ADM65290/c  
ID ADM65290 standard; DNA; 289 BP.

AC ADM65290;  
XX 03-JUN-2004 (first entry)  
DT  
XX

DE NRY polymorphism detection primer #299.

XX ethnic origin determination; polymorphic site determination;

KM Y chromosome; paternity testing; forensic; diagnosis;

XX non-recombining region; human; NRY; PCR; primer; ss.

OS Homo sapiens.

XX US2003134285-A1.

XX 17-JUL-2003.

XX 01-NOV-2001; 2001US-00002623.

XX 01-NOV-2000; 2000US-0245355P.

XX (OEFPN/) OEFNER P J.

XX (UNDE/) UNDERHILL P A.

XX Oefner PJ, Underhill PA.

XX WPI; 2003-843259/78.

XX Determining the ethnic origin of a male by obtaining a nucleic acid

PT sample from the male and identifying at least two polymorphic markers in

PT the nucleic acid sample indicative of the ethnic origin of the male.

XX Claim 24; Page 39; 74pp; English.

XX The invention describes a method of determining the ethnic origin of a

CC male comprising obtaining a nucleic acid sample from the male, and

CC identifying at least two polymorphic markers in the nucleic acid sample

CC indicative of the ethnic origin of the male, using at least one primer

CC pair from the primer pairs given in the specification. Also described is

CC a method of: identifying polymorphic sites in a nucleic acid; a kit for

CC determining the ethnic origin of an individual; determining the ethnic

CC origin of a human male individual; an isolated nucleic acid segment of a

CC human Y chromosome comprising at least 10 contiguous bases including at

CC least one of the polymorphic sites given in the specification; nucleic

CC acid primer pairs for amplifying polymorphic regions of the Y chromosome

CC given in the specification; and determining the paternity of a human male

CC individual. The method is useful for determining the ethnic origin of a

CC male, for paternity testing, for forensic studies or for diagnosis. This

CC sequence represents a primer used to detect polymorphisms in the non-

CC recombining region of the human Y chromosome (NRY).

XX Sequence 289 BP; 71 A; 76 C; 89 G; 52 T; 0 U; 1 Other;

QY Query Match 88.9%; Score 16; DB 11; Length 289;

AC Best Local Similarity 100.0%; Pred. No. 5.2e+02;

XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 281 CCCCCCTTTGTGATA 266

RESULT 8  
ACF84919  
ID ACF84919 standard; DNA; 448 BP.

AC ACF84919;  
XX 02-JUN-2005 (first entry)  
DT  
XX

DE Human SIRS/sepsis diagnostic marker DNA fragment 3779.

XX Systemic inflammatory response syndrome; SIRS; antibacterial;

KM immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.

OS Homo sapiens.

XX WO2004087949-A2.



PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 XX (BLIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Hanelbeck R, Ohlsen KL, Zyekind JW,  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 P-PSDB; ABU23606.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 14; SEQ ID NO 15346; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1104 BP; 443 A; 95 C; 239 G; 327 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 8; Length 1104;

Best Local Similarity 94.1%; Pred. No. 1.1e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCTCTCTTGGATAC 18  
 |||||  
 DB 865 CCCCTCTTGGATAC 849

RESULT 11  
 AAQ04292/c  
 ID AAQ04292 standard; DNA; 1866 BP.  
 XX

AC AAQ04292;

XX 25-MAR-2003 (revised)  
 DT 24-SEP-1990 (first entry)  
 XX

DE Encodes polypeptide recognised by neutralising anti-HSV antibodies.  
 XX

KW Fish haemorrhagic septicemia virus (HSV);  
 KW immunogenic recombinant polypeptide; ss.

OS Synthetic.

XX CA2000570-A.

XX 12-APR-1990.

XX 10-DEC-1989; 89CA-02000570.

XX 12-OCT-1988; 88EP-00402587.

XX (RENA/) RENARD A J J.

XX (EURO-) EUROGENTEC SA.

XX (BURO-) BURGENTEC SA.

XX Renard AJJ, Thiry MEG;

XX WPI; 1990-140005/19.

XX P-PSDB; AAR04576.

XX Immunogenic recombinant polypeptide - from fish haemorrhagic septicemia  
 XX virus for protective vaccines, and new DNA sequences, vectors,  
 XX transformed hosts and antibodies.

XX Claim 13; Page 72; 99pp; French.

XX Recombinant polypeptide encoded by this sequence is included in all

CC recombinant proteins covered by the invention. See also AAQ04308,  
 CC AAQ04318-20, AAQ05332-42. (Updated on 25-MAR-2003 to correct PF field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 1866 BP; 536 A; 503 C; 434 G; 393 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 2; Length 1866;

Best Local Similarity 94.1%; Pred. No. 1.2e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCTCTCTTGGATA 17  
 |||||

DB 1384 CCCCTCTTGGATA 1368

RESULT 12

AD531146/c  
 ID AD531146 standard; DNA; 6388 BP.

XX AD531146;

XX 18-NOV-2004 (first entry)

XX Human genome high complexity repeat found in the HIRA gene #179.

XX Human; ds;  
 KW histone cell cycle regulation defective, *S. cerevisiae* homologue A; HIRA;  
 KW high complexity repeat; in situ hybridisation; Southern blot;  
 KW chromosome breakpoint; inherited genetic disease; neoplastic disorder;  
 KW chromosome 22; DiGeorge syndrome; Velo-Cardio-facial syndrome.

XX Homo sapiens.

XX US2003224356-A1.

XX 04-DEC-2003.

XX 14-MAY-2001; 2001US-00854867.

XX 16-MAY-2000; 2000US-00573080.

XX (KNOL/) KNOLL J H M.

XX (ROGA/) ROGAN P K.

PI Knoll JHM, Rogan PK;  
XX WPI; 2002-062378/08.  
XX  
XX  
PT Single copy genomic hybridization probes for detecting specific nucleic  
PT acid sequences in sample by in situ hybridization useful for detection of  
PT acquired or inherited genetic diseases.

XX Example 1; SEQ ID NO 179; 30pp; English.

XX The invention relates to a nucleic acid hybridization probe comprising a  
XX labelled, single copy nucleic acid of at least 50 nucleotides, which  
XX will hybridize to a deduced single copy sequence interval in target  
XX nucleic acid (TNA) of known sequence. The single copy sequence is deduced  
XX by comparing the target nucleic acid (e.g. a disease causing gene) with a  
XX collection of high and low complexity repeat sequences as found in the  
XX genome of the organism from containing the target nucleic acid. The probe  
XX is generated by PCR on the target sequence. The probe is essentially free  
XX of blocking nucleic acid sequences which will hybridize to repeat  
XX sequences within the genome of which the TNA is a part, and is labelled  
XX with a label selected from fluorochrome-responsive labels, fluorochromes,  
XX calorimetric chemical, conjugated proteins, antibodies, antigens and  
XX their mixtures. The probe is useful in a hybridization method, where the  
XX hybridization method is from in situ hybridization, Southern blot, and  
XX other methods in which nucleic acid is immobilized, where the method  
XX further comprises selecting a single copy nucleic acid which will  
XX hybridize to a duplication or triplication sequence domain. The probe is useful  
XX for determining the existence of previously unknown repeat sequence  
XX families in a genome. The method comprises reacting a labelled probe with  
XX the genome, causing the probe to hybridize and ascertaining if the probe  
XX hybridizes to the genome at more than three preferably ten different  
XX locations as a determination of new repeat sequence family, where the  
XX determining step comprises selecting the single copy sequence from a  
XX duplication or triplication sequence domain. The probe is useful for  
XX determining a chromosome breakpoint and is useful in the fields for  
XX cytogenetics and molecular genetics for determining the presence of  
XX specific nucleic acid sequences in a sample of eukaryotic origin, e.g.  
XX the probes may be used to analyse specific chromosomal locations by in  
XX situ hybridization as a detection of acquired or inherited genetic  
XX diseases especially for detection of genetic or neoplastic disorders.  
XX Unlike prior art techniques, the probe permits more precise chromosomal  
XX breakpoint determinations by in situ hybridization. The genomic sequence  
XX comprising the human HIRA gene (histone cell cycle regulation defective,  
XX S. cerevisiae, homologue A) was analysed for single copy sequence  
XX intervals for use as probes of the invention. HIRA is located on  
XX chromosome 22 as a duplicate, deletions of 1 copy lead to DiGeorge and  
XX Velo-Cardio-facial syndromes. The present sequence is a high complexity  
XX repeat found within the human genome used to analyse the HIRA gene for  
XX repeat regions. Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic format  
XX directly from USPTO at seqdata.uspto.gov/sequence.html?docID=20030224356.

XX Sequence 6388 BP; 2006 A; 1007 C; 1040 G; 2211 T; 0 U; 124 Other;

XX Query Match 85.6%; Score 15.4; DB 7; Length 6388;  
XX Best Local Similarity 94.1%; Pred. No. 1.3e+03;  
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGATA 17  
DB 4010 CCCACCTCTCTGATA 3994

RESULT 13  
ADY36534/C  
ID ADY36534 standard; DNA; 6388 BP.

AC ADY36534;

DT 05-MAY-2005 (first entry)

DE HIRA genomic fragment SEQ ID NO 179.

XX

KW hybridization; DNA detection; neoplasm; genetic disorder; cytogenetics;  
KW HIRA, ds.

OS Homo sapiens.

XX MO200188089-A2.

XX 22-NOV-2001.

XX 15-MAY-2001; 2001WO-US015674.

XX 16-MAY-2000; 2000US-00573080.

XX 14-MAY-2001; 2001US-00854867.

XX (CHIL-) CHILDREN'S MERCY HOSPITAL.

PI Knoll JHM, Rogan PK, Cazarro PM;

XX WPI; 2002-062378/08.

XX Single copy genomic hybridization probes for detecting specific nucleic  
XX acid sequences in sample by in situ hybridization useful for detection of  
XX acquired or inherited genetic diseases.

XX Example 1; SEQ ID NO 179; 67pp; English.

XX The invention describes a nucleic acid hybridization probe (I) comprising  
XX a labeled, single copy nucleic acid of at least 50 nucleotides, which  
XX will hybridize to a deduced single copy sequence interval in target  
XX nucleic acid (TNA) of known sequence. (I) is useful in a hybridization  
XX method which comprises preparing a reaction mixture comprising TNA and  
XX (I) which hybridizes to TNA, and causing (I) to hybridize to TNA, where  
XX the hybridization method is from in situ hybridization, Southern blot,  
XX and other methods in which nucleic acid is immobilized, where the method  
XX further comprises selecting a single copy nucleic acid which will  
XX hybridize to a duplication or triplication sequence domain. (I) is useful for:  
XX determining the existence of previously unknown repeat sequence families  
XX in a genome; determining a chromosome breakpoint and in the fields of  
XX cytogenetics and molecular genetics for determining the presence of  
XX specific nucleic acid sequences in a sample of eukaryotic origin, e.g.  
XX the probes may be used to analyse specific chromosomal locations by in  
XX situ hybridization as a detection of acquired or inherited genetic  
XX diseases especially for detection of genetic or neoplastic disorders.  
XX Unlike prior art techniques, (I) permits more precise chromosomal  
XX breakpoint determinations by in situ hybridization. Hybridization  
XX techniques utilizing (I), have made it possible to obtain reliable,  
XX easily detectable signals with relatively small probes. A readily  
XX detectable signal was obtained with a probe on the order of 2 kb in  
XX length, using fluorescent in situ hybridization (FISH) technology. This  
XX sensitivity of (I) is improved compared to the prior art, because the  
XX probes of (I) are homogeneous single copy sequences. However, smaller  
XX amplified segments, each comprising non-repetitive sequences, may also be  
XX used in combination as probes to achieve adequate signals for in situ  
XX hybridization. Complex single copy probes that hybridize to duplicated or  
XX triplicated targets can also increase hybridization signals. This  
XX sequence represents a human HIRA genomic sequence that shows homology to  
XX a known high-complexity repeat sequence family of the human genome and is  
XX used in the creation of an HIRA gene probe.

XX Sequence 6388 BP; 2006 A; 1007 C; 1040 G; 2211 T; 0 U; 124 Other;

XX Query Match 85.6%; Score 15.4; DB 7; Length 6388;  
XX Best Local Similarity 94.1%; Pred. No. 1.3e+03;  
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGATA 17  
DB 4010 CCCACCTCTCTGATA 3994

RESULT 14  
ABU12362  
ID ABU12362 standard; cDNA; 12542 BP.

```
XX ABL12362;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 31568.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
DR
XX
XX P-PsDB; ABB68259.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 31568; 21bp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 12542 BP; 3522 A; 2666 C; 2797 G; 3557 T; 0 U; 0 Other;
XX
XX Query Match 85.6%; Score 15.4; DB 4; Length 12542;
XX Best Local Similarity 94.1%; Pred. No. 1.4e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 CCCCCCTCTTCGATA 17
XX |||||
XX DB 5329 CCCCCCTCTTCGATA 5345
XX
XX RESULT 15
XX ACN45006
XX ID ACN45006 standard; DNA; 67810 BP.
XX
XX ACN45006;
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX Human genomic sequence hCG32956.
DE
XX
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2003073826-A2.
XX
XX 12-SEP-2003.
XX
```

```
XX
XX 28-FEB-2003; 2003WO-US006235.
PF
XX
XX 01-MAR-2002; 2002US-00087192.
PR
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
XX
XX WPI; 2003-328604/31.
DR
XX
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 1738; opp; English.
PS
XX
XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
XX SQ Sequence 67810 BP; 19583 A; 15025 C; 13900 G; 19302 T; 0 U; 0 Other;
XX
XX Query Match 85.6%; Score 15.4; DB 11; Length 67810;
XX Best Local Similarity 94.1%; Pred. No. 1.6e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 CCCCCCTCTTCGATA 17
XX |||||
XX DB 2335 CCCCCCTCTTCGATA 23369
XX
XX RESULT 16
XX ABD33177/C
XX ID ABD33177 standard; DNA; 79256 BP.
XX
XX ABD33177;
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX Murine cancer-associated (CA) gene MD07-026.
DE
XX
XX Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
KM ds; cancer; cytostatic.
XX
XX Mus musculus.
XX
XX WO2004058146-A2.
XX
XX 15-JUL-2004.
XX
XX 15-DEC-2003; 2003WO-US040081.
XX
XX 17-DEC-2002; 2002US-00322281.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
XX
XX WPI; 2004-499109/47.
XX
XX Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX
```



```
XX PS Disclosure; SEQ ID NO 167; 182pp; English.
XX CC
XX CC The invention relates to cancer-associated proteins (CAP) and the cancer-
XX CC associated (CA) nucleic acids encoding them. The invention also relates
XX CC to a method for treating cancers involving administering to a patient an
XX CC inhibitor of CAP, and a method of screening for anticancer activity in a
XX CC potential drug involving providing a cell that expresses a CA gene,
XX CC contacting a tissue sample derived from a cancer cell with an anticancer
XX CC drug candidate and monitoring the effect of the anticancer drug candidate
XX CC on expression of the CA gene. The CAP proteins are useful for detecting
XX CC cancer associated with expression of a CAP protein in a test cell sample
XX CC and for screening for a bioactive agent capable of modulating the
XX CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
XX CC cancer, involving determining the expression of a CA nucleic acid in a
XX CC tissue. This sequence represents a murine CA gene of the invention. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX Sequence 79256 BP; 19839 A; 19064 C; 19416 G; 20480 T; 0 U; 457 Other;
XX
XX Query Match 85.6%; Score 15.4; DB 13; Length 79256;
XX Best Local Similarity 94.1%; Pred. No. 1.6e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 CCCCCCTCTCTGGATA 17
XX Db 35784 CCCCCCTCTCTGGATA 35768
XX
XX RESULT 17
XX ADH10017.1/c
XX Continuation (2 of 4) of ADH10017 from base 100001 (Human chromosome 2p21-22 fragment cd
XX WP Sequence split into 4 fragments LOCUS ADH10017 Accession Adh10017
XX WP Fragment Name Begin End
XX WP ADH10017_0 1 110000
XX WP ADH10017_1 100001 210000
XX WP ADH10017_2 200001 310000
XX WP ADH10017_3 300001 365186
XX
XX Query Match 85.6%; Score 15.4; DB 10; Length 110000;
XX Best Local Similarity 94.1%; Pred. No. 1.6e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 2 CCCCCCTCTCTGGATAC 18
XX Db 22603 CACCCCTCTCTGGATAC 22587
XX
XX RESULT 18
XX ADQ97266.o/c
XX WP Sequence split into 4 fragments LOCUS ADQ97266 Accession Adq97266
XX WP Fragment Name Begin End
XX WP ADQ97266_0 1 110000
XX WP ADQ97266_1 100001 210000
XX WP ADQ97266_2 200001 310000
XX WP ADQ97266_3 300001 355211
XX ID ADQ97266 standard; DNA; 355211 BP.
XX
XX ADQ97266;
XX AC
XX AC
XX DT 07-OCT-2004 (first entry)
XX DB Human cancer associated sequence HD08-023, SEQ ID 242.
XX KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX OS Homo sapiens.
XX XX WO2004060304-A2.
XX FN
XX PD 22-JUL-2004.
XX
XX XX
XX PF 22-DEC-2003; 2003MO-US041389.
XX XX
XX PR 27-DEC-2002; 2002US-00330773.
XX XX
XX PA (SAGR-) SAGRES DISCOVERY INC.
XX XX
XX PI Morris DW, Malandro MS;
XX DR
XX DR WPI; 2004-543781/52.
XX XX
XX PT New isolated cancer associated nucleic acids comprising at least 10
XX PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX PT cancers such as leukemia and lymphoma.
XX PS
XX PS Claim 1; SEQ ID NO 242; 199pp; English.
XX XX
XX XX The present invention relates to cancer associated sequences (ADQ97025-
XX CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX CC data for this patent did not form part of the printed specification, but
XX CC was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ
XX Sequence 355211 BP; 96742A; 68188C; 72417G; 106131T; 0U; 117330Other;
XX
XX Query Match 85.6%; Score 15.4; DB 12; Length 110000;
XX Best Local Similarity 94.1%; Pred. No. 1.6e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 2 CCCCCCTCTCTGGATAC 18
XX Db 102906 CACCCCTCTCTGGATAC 102890
XX
XX RESULT 19
XX ADQ97266.1/c
XX Continuation (2 of 4) of ADQ97266 from base 100001 (Human cancer associated sequence HD
XX WP Sequence split into 4 fragments LOCUS ADQ97266 Accession Adq97266
XX WP Fragment Name Begin End
XX WP ADQ97266_0 1 110000
XX WP ADQ97266_1 100001 210000
XX WP ADQ97266_2 200001 310000
XX WP ADQ97266_3 300001 355211
XX
XX Query Match 85.6%; Score 15.4; DB 12; Length 110000;
XX Best Local Similarity 94.1%; Pred. No. 1.6e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 2 CCCCCCTCTCTGGATAC 18
XX Db 2906 CACCCCTCTCTGGATAC 2890
XX
XX RESULT 20
XX ACH28985/c
XX ID ACH28985 standard; cDNA; 409 BP.
XX AC
XX AC ACH28985;
XX AC
XX DT 13-OCT-2003 (first entry)
XX DB Human adult spleen cDNA #4.
XX DE
XX DE Human adult spleen cDNA #4.
XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX KW genome mapping; biodiversity; genetic disorder.
XX OS Homo sapiens.
XX XX US2003073623-A1.
XX FN
XX PD 17-APR-2003.
XX XX
```

PF 30-JUL-2001; 2001US-00918995.  
XX  
XX  
PR 30-JUL-2001; 2001US-00918995.  
XX  
XX  
PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
XX  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX  
XX  
DR MPI; 2003-615964/58.  
XX  
XX  
PT New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX  
PS Claim 1, SEQ ID NO 16197; 44pp; English.  
XX  
XX  
CC The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
CC  
XX  
SQ Sequence 409 BP; 123 A; 92 C; 110 G; 83 T; 0 U; 1 Other;  
XX  
XX  
Query Match 83.3%; Score 15; DB 9; Length 409;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 2 CCCCTCTCTTGAT 16  
Db 168 CCCCTCTCTTGAT 154  
XX  
XX  
RESULT 21  
ADG31675/c  
ID ADG31675 standard; cDNA; 2200 BP.  
XX  
XX  
AC ADG31675;  
XX  
XX  
DT 26-FEB-2004 (first entry)  
XX  
XX  
DE Human cDNA encoding a protein with VEGF promoting activity SeqID21.  
XX  
XX  
KW human; gene; ss; vascularisation; vascular endothelial growth factor;  
KW VEGF; tumour metastasis; psoriasis; chronic rheumatoid arthritis;  
KW diabetic retinopathy; angiodioma; macular degeneration;  
KW angina pectoris; cerebral infarction; vascular dementia;  
KW Burger's disease; antiinflammatory; cytostatic; antiarthritic;  
KW antiproliferative; antiarteriosclerotic; cardiant; vasotropic;  
KW neuroprotective.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO2003083117-A2.  
XX  
XX  
PD 09-OCT-2003.

XX  
XX  
PF 01-APR-2003; 2003WO-JP004163.  
XX  
XX  
PR 01-APR-2002; 2002JP-0008679.  
PR 02-APR-2002; 2002US-0368977P.  
XX  
XX  
PA (ASAH ) ASAMI KASEI KK.  
XX  
XX  
PI Nagumo H, Matsuda A, Sugahara T;  
XX  
XX  
DR MPI; 2003-812544/76.  
DR P-PSDB; ADG31676.  
XX  
XX  
PT New purified protein, useful for preparing a composition for treating or  
PT preventing a disease associated with abnormal expression of VEGF or VEGF  
PT receptor, e.g., solid tumor, inflammation, psoriasis or chronic  
PT rheumatoid arthritis.  
XX  
XX  
PS Claim 4, SEQ ID NO 21; 405pp; English.  
XX  
XX  
CC This invention relates to novel proteins having vascularisation activity.  
CC Specifically, it refers to proteins having an activity of promoting the  
CC expression of vascular endothelial growth factor (VEGF) and/or the VEGF  
CC receptor. VEGF is a representative growth factor involved in normal  
CC vascularisation in growing tissues, for example for fetal growth, tissue  
CC repair and for the menstrual cycle and pregnancy. The present invention  
CC describes these purified proteins, vectors and appropriate antibodies, as  
CC well as a screening assay for identifying substances capable of  
CC inhibiting or promoting VEGF (and VEGF receptor) expression. Furthermore,  
CC it refers to compositions useful for the diagnosis, treatment and  
CC prevention of diseases associated with abnormal VEGF/ VEGF receptor  
CC expression, such as tumor metastasis, psoriasis, chronic rheumatoid  
CC arthritis, diabetic retinopathy, angiodioma, macular degeneration,  
CC angina pectoris, cerebral infarction, vascular dementia or Burger's  
CC disease. Accordingly, they exhibit antiinflammatory, cytostatic,  
CC antiarthritic, antiproliferative, antiarteriosclerotic, cardiant, vasotropic  
CC and neuroprotective activities. This polynucleotide sequence is a human  
CC cDNA that encodes a protein with VEGF promoting activity of the  
CC invention.  
XX  
XX  
SQ Sequence 2200 BP; 666 A; 423 C; 541 G; 570 T; 0 U; 0 Other;  
XX  
XX  
Query Match 83.3%; Score 15; DB 10; Length 2200;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 2 CCCCTCTCTTGAT 16  
Db 500 CCCCTCTCTTGAT 486  
XX  
XX  
RESULT 22  
AAD02351/c  
ID AAD02351 standard; cDNA; 2405 BP.  
XX  
XX  
AC AAD02351;  
XX  
XX  
DT 24-APR-2001 (first entry)  
XX  
XX  
DE Human RNA metabolism protein (RMEP-8) cDNA.  
XX  
XX  
KW Human; RNA metabolism protein; RMEP; neurotrophic; neuroleptic; antifungal;  
KW tranquilliser; antianaemic; antidiabetic; immunosuppressive; cytostatic;  
KW antiasthmatic; antiinflammatory; anti-HIV; human immunodeficiency virus;  
KW antiautistic; antiarteriosclerotic; antiatherosclerotic; antiallergic;  
KW antirheumatic; antiparkinsonian; antithyroid; nephrotoxic; antigout;  
KW thymomimetic; RMEP expression modulator; transgenic; spinal cord disease;  
KW nervous system disorder; Alzheimer's disease; therapy; gene therapy;  
KW neuromuscular disorder; hepatitis; cancer; cell proliferative disorder;  
KW peripheral nervous system disorder; cirrhosis; cranial nerve disorder;  
KW progressive neural autonomic nervous system disorder; Addison's disease;  
KW amyotrophic lateral sclerosis; autoimmune disorder; drug screening; ss.  
XX



RESULT 24  
ADG31677/C  
ID ADG31677 standard; cDNA; 2856 BP.  
XX  
AC ADG31677;  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Human cDNA encoding a protein with VEGF promoting activity SeqID23.  
XX  
KW human; gene; ss; vascularisation; vascular endothelial growth factor;  
KW VEGF; tumor metastasis; psoriasis; chronic rheumatoid arthritis;  
KW diabetic retinopathy; angiodysplasia; macular degeneration;  
KW angina pectoris; cerebral infarction; vascular dementia;  
KW Buerger's disease; antiphospholipid; cytosolic; antiarthritic;  
KW antipsoriatic; antiarteriosclerotic; cardiac; vasotropic;  
KW neuroprotective.  
XX  
OS Homo sapiens.  
XX  
PN WO2003083117-A2.  
XX  
PD 09-OCT-2003.  
XX  
PP 01-APR-2003; 2003WO-JP004163.  
XX  
PR 01-APR-2002; 2002JP-00098679.  
XX  
PR 02-APR-2002; 2002US-0368977P.  
XX  
PA (ASAH ) ASAMI KASEI KK.  
XX  
PI Nagumo H, Matsuda A, Sugahara T;  
XX  
DR WPI; 2003-812544/76.  
XX  
DR P-PSDB; ADG31678.  
XX  
PT New purified protein, useful for preparing a composition for treating or  
PT preventing a disease associated with abnormal expression of VEGF or VEGF  
PT receptor, e.g., solid tumor, inflammation, psoriasis or chronic  
PT rheumatoid arthritis.  
XX  
PS Claim 4; SEQ ID NO 23; 405bp; English.  
XX  
CC This invention relates to novel proteins having vascularisation activity.  
CC Specifically, it refers to proteins having an activity of promoting the  
CC expression of vascular endothelial growth factor (VEGF) and/or the VEGF  
CC receptor. VEGF is a representative growth factor involved in normal  
CC vascularisation in growing tissues, for example for fetal growth, tissue  
CC repair and for the menstrual cycle and pregnancy. The present invention  
CC describes these purified proteins, vectors and appropriate antibodies, as  
CC well as a screening assay for identifying substances capable of  
CC inhibiting or promoting VEGF (and VEGF receptor) expression. Furthermore,  
CC it refers to compositions useful for the diagnosis, treatment and  
CC prevention of diseases associated with abnormal VEGF/VEGF receptor  
CC expression, such as tumor metastasis, psoriasis, chronic rheumatoid  
CC arthritis, diabetic retinopathy, angiodysplasia, macular degeneration,  
CC angina pectoris, cerebral infarction, vascular dementia or Buerger's  
CC disease. Accordingly, they exhibit antiinflammatory, cytosolic,  
CC antiphospholipid, antiphospholipid, antiarteriosclerotic, cardiac, vasotropic  
CC and neuroprotective activities. This polynucleotide sequence is a human  
CC cDNA that encodes a protein with VEGF promoting activity of the  
CC invention.  
XX  
SQ Sequence 2856 BP; 862 A; 594 C; 689 G; 711 T; 0 U; 0 Other;  
XX  
Query Match 83.3%; Score 15; DB 10; Length 2856;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 CCCCTCTTCTGGAT 16  
DB 1156 CCCCTCTTCTGGAT 1142

RESULT 25  
ACN39725/C  
ID ACN39725 standard; cDNA; 5208 BP.  
XX  
AC ACN39725;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Tumour-associated antigenic target (TAT) cDNA DNA325917, SEQ ID NO:4050.  
XX  
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
KW tumor; diagnosis; cell proliferative disorder; breast cancer;  
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
KW central nervous system cancer; bladder cancer; pancreatic cancer;  
KW cervical cancer; melanoma; leukemia; hybridisation probe;  
KW chromosome identification; chromosome mapping; gene mapping;  
KW gene therapy; cytosolic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2004030615-A2.  
XX  
PD 15-APR-2004.  
XX  
PP 29-SEP-2003; 2003WO-US028547.  
XX  
PR 02-OCT-2002; 2002US-0414971P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Wu TD, Zhang Z, Zhou Y;  
XX  
DR WPI; 2004-347921/32.  
XX  
DR P-PSDB; ABW81569.  
XX  
PT New tumor-associated antigenic target polypeptides and nucleic acids,  
PT useful in preparing a medicament for treating or detecting a  
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
PT prostate cancer or tumor.  
XX  
PS Claim 1; SEQ ID NO 4050; 7273bp; English.  
XX  
CC The invention relates to human tumour-associated antigenic target (TAT)  
CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
CC overexpressed in cancer tissues compared to normal tissues, and may thus  
CC serve as effective targets for the diagnosis and treatment of cancer in  
CC mammals. The invention also relates to nucleic acid and polypeptide  
CC sequences at least 80% identical to the TAT nucleic acids and  
CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
CC TAT polypeptide; and methods and compositions for the treatment or  
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
CC antibodies, antagonists, binding molecules and compositions are useful  
CC for diagnosing or treating a cell proliferative disorder associated with  
CC increased TAT expression, particularly cancers such as breast cancer,  
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT nucleic acid of the invention  
XX  
SQ Sequence 5208 BP; 1607 A; 1105 C; 1286 G; 1210 T; 0 U; 0 Other;  
XX  
Query Match 83.3%; Score 15; DB 13; Length 5208;  
Best Local Similarity 100.0%; Pred. No. 2e+03;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 CCCCTCTTCTGGAT 16  
DB 3401 CCCCTCTTCTGGAT 3387

ID	ADK65780	standard; DNA; 5581 BP.
AC	ADK65780;	
XX		
D7	06-MAY-2004	(first entry)
XX		
DE	Angiogenesis-differentially expressed gene ANH0144B.	
XX		
KW	ds; gene; cytoskeletal; cardiac; vasotropic; antiarteriosclerotic;	
KW	angiotensin II receptor; angiotensin II receptor agonist; angiogenic index;	
KW	protein expression; cancer; coronary artery disease; myocardial ischemia;	
KW	coronary arteriosclerosis; forensic medicine.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	264..333
FT		/tag= a
FT		/product= "Angiogenesis protein ANH0144B"
PX	WO2003066831-A2.	
PD	14-AUG-2003.	
XX		
PF	07-FEB-2003; 2003WO-US003848.	
XX		
PR	07-FEB-2002; 2002US-0067482.	
PR	10-JUN-2002; 2002US-00164595.	
PR	16-AUG-2002; 2002US-0403649P.	
PR	03-JAN-2003; 2003US-0437746P.	
PA	(ORIG-) ORIGENE TECHNOLOGIES INC.	
PI	Sun Z, Li X, Kovacs KF, Fan W, Jay G;	
DR	WPI; 2003-731502/69.	
PT	P-PADB; ADK65781.	
PT	Determining the angiogenic index of a tissue or cell sample using	
PT	expression levels of differentially expressed genes, useful for	
PT	diagnosing or treating cancer, coronary artery disease, myocardial	
PT	ischemia and/or arteriosclerosis.	
PS	Claim 19; SEQ ID NO 19; 296pp; English.	
CC	The invention relates to a method of determining the angiogenic index of	
CC	a tissue or cell sample comprising assessing, in a sample, the expression	
CC	levels of one or more differentially-expressed protein from any of 34 DNA	
CC	sequences, given in the specification, where the levels are indicative of	
CC	the angiogenic index. The methods and compositions of the present	
CC	invention are useful for diagnosing, preventing and/or treating cancer,	
CC	coronary artery disease, myocardial ischemia or coronary	
CC	arteriosclerosis. They can also be used in research, drug discovery and	
CC	forensic medicine involving angiotensins. This sequence corresponds to	
CC	one of the differentially expressed genes of the invention.	
XX		
SEQ	Sequence 5581 BP; 1683 A; 1280 C; 1200 G; 1418 T; 0 U; 0 Other;	
Query Match	83.3%; Score 15; DB 10; Length 5581;	
Best Local Similarity	100.0%; Pred. No. 2.le+03;	
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0.		
OY	2 CCCCTCTTCTGGAT 16	
DB	2348 CCCCTCTTCTGGAT 2362	

ID	AB061143	standard; cDNA; 5647 BP.
XX		
AC	AB061143;	
XX		
DT	26-FEB-2003	(first entry)
XX		
DE	Nuclear protein Zap encoding sequence.	
XX		
KM	Neuroprotective; immunomodulator; cancer; chromosome 14q24.3; cytostatic;	
XX	anti-inflammatory; gene therapy; nutritional supplement; wound; burn;	
KW	ulcer; Alzheimer's disease; Huntington's disease;	
XX	amyotrophic lateral sclerosis; autoimmune disorder; inflammation;	
KW	vulnerary; gene; ss.	
XX		
OS	Mus musculus.	
XX		
PN	WO200231111-A2.	
XX		
PD	18-APR-2002.	
XX		
PF	11-OCT-2001; 2001WO-US027760.	
XX		
PR	12-OCT-2000; 2000US-00687527.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang Y <sup>T</sup> , Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;	
XX	Xue AJ, Yang Y, Wehrman T, Drmanac RT;	
XX		
DR	WPI; 2002-426278/45.	
XX	N-PSDB; ABP43899.	
PT	New polypeptides and their encoded proteins, useful as nutritional	
XX	sources or supplements, or in gene therapy, particularly for treating	
PT	wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or	
XX	inflammation.	
PS	Claim 1; SEQ ID # 356; 357bp + Sequence Listing; English.	
XX		
CC	The invention relates to 446 newly isolated polynucleotide sequences. The	
XX	activity of polynucleotides of the invention may be described as,	
CC	vulnerary, neuroprotective, immunomodulator, cytostatic and anti-	
XX	inflammatory. Compositions comprising nucleic acids of the invention are	
CC	useful for treating a mammalian subject, or as nutritional sources or	
XX	supplements. There are useful in gene therapy, particularly for treating	
CC	wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,	
XX	amyotrophic lateral sclerosis, autoimmune disorders, cancer or	
CC	inflammation. The nucleic acids and polypeptides are also useful in	
XX	diagnostic and research methods. The sequences given in records AB060788-	
CC	AB061233 represent polynucleotides of the invention. NOTE: The sequence	
XX	data for this patent did not form part of the printed specification, but	
CC	was obtained in electronic form directly from WIPO at	
XX	ftp.wipo.int/pub/published_pmt_sequences	
XX		
XX	Sequence 5647 BP; 1603 A; 1473 C; 1418 G; 1153 T; 0 U; 0 Other;	
QY		
2	CCCCCTCTTCGGAT 16	
Best Local Similarity	83.3%; Score 15; DB 6; Length 5647;	
Matches	15; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
DB	4872 CCCCCTCTTCGGAT 4858	
RESULT 28		
ID	AD30216/C	
XX	AD30216 standard; cDNA; 6367 BP.	
XX		
AC	AD30216;	
XX		
DT	18-DEC-2003	(first entry)
XX		

DE Human novel cDNA sequence, SEQ ID NO:298.  
XX  
KW Human; diagnostic; drug screening; forensics; gene mapping;  
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
KW ulcers; osteoporosis; autoimmune disease; cancer;  
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;  
KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;  
KW anticancer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
KW gene therapy; chromosome 14q24.3; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO2003029271-A2.  
XX  
XX 10-APR-2003.  
XX  
XX 24-SEP-2002; 2002WO-US030474.  
XX  
XX 24-SEP-2001; 2001US-0324631P.  
XX  
XX (HYSB-) HYSEQ INC.  
XX  
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Mehrman T;  
PI Zhou P, Ghosh M, Mang D, Ma Y, Asundi V, Wang Z, Weng G;  
PI Haley-Vicente D, Drmanac RT;  
XX  
XX WPI; 2003-371981/35.  
XX  
XX P-PSDB; ADC31187.  
XX  
XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anemias, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.  
XX  
XX Claim 1; SEQ ID NO 298; 1185bp; English.  
XX  
XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 9% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC1861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a specifically  
CC claimed human cDNA sequence of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 6367 BP; 1793 A; 1695 C; 1555 G; 1324 T; 0 U; 0 Other;

Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 CCCCTCTTCGTGAT 16  
DB 5472 CCCCTCTTCGTGAT 5458  
RESULT 29  
ID ADC85744/C  
ID ADC85744 standard; DNA; 60989 BP.  
XX  
AC ADC85744;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Human GPCR gene SEQ ID NO:197.  
XX  
KW ds; gene; human; GPCR;  
KW guanosine triphosphate-binding protein coupled receptor; gene therapy.  
XX  
XX Homo sapiens.  
XX  
XX EP1270724-A2.  
XX  
XX 02-JAN-2003.  
XX  
XX 18-JUN-2002; 2002EP-00013517.  
XX  
XX 18-JUN-2001; 2001JP-00246789.  
XX  
XX (NABD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX (ABSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.  
XX  
XX Suwa M, Asai K, Akiyama Y, Aburatani H;  
PI WPI; 2003-315783/31.  
XX  
XX P-PSDB; ADC85745.  
XX  
XX New polynucleotide, useful for preparing a composition for treating a  
PT patient in need of increased or suppressed activity or expression of the  
XX guanosine triphosphate-binding protein coupled receptor.  
XX  
XX Claim 1; SEQ ID NO 197; 28bp; English.  
XX  
XX The invention relates to a novel polynucleotide encoding a guanosine  
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of  
CC the invention may have a use in gene therapy. The polynucleotide and  
CC polypeptide are useful for preparing a composition for treating a patient  
CC in need of increased or suppressed activity or expression of the  
CC guanosine triphosphate-binding protein coupled receptor. The  
CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the  
CC invention.  
XX  
SQ Sequence 60989 BP; 18566 A; 11696 C; 12057 G; 1766 T; 0 U; 904 Other;

Query Match 83.3%; Score 15; DB 10; Length 60989;  
Best Local Similarity 100.0%; Pred. No. 2.5e+03;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 CCCCTCTTCGTGATAC 18  
DB 51268 CCCCTCTTCGTGATAC 51254  
RESULT 30  
ABD32627 0  
WP Sequence split into 5 fragments LOCUS ABD32627 Accession Abd32627  
WP Fragment Name Begin End  
WP ABD32627\_0 1 110000  
WP ABD32627\_1 100001 210000  
WP ABD32627\_2 200001 310000  
WP ABD32627\_3 300001 410000

WP ABD32627\_4 40001 421609  
ID ABD32627 standard; DNA; 421609 BP.  
XX  
AC ABD32627;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Mouse cancer-associated genomic DNA MD13-017.  
XX  
KM Mouse; da; cancer-associated protein; gene; cytostatic; cancer;  
XX leukemia; lymphoma; CAP.  
XX  
OS Mus musculus.  
XX  
PN MO2004074320-A2.  
XX  
PD 02-SEP-2004.  
XX  
PF 17-FEB-2004; 2004MO-US004730.  
XX  
PR 14-FEB-2003; 2003US-00367094.  
PR 14-MAR-2003; 2003US-00388838.  
PR 15-APR-2003; 2003US-00417375.  
PR 13-JUN-2003; 2003US-00461862.  
PR 15-SEP-2003; 2003US-00663431.  
PR 15-DEC-2003; 2003US-00737318.  
XX  
PA (SAGR-) SAGRES DISCOVERY INC.  
XX  
PI Morris DW, Morris DW, Malandro MS;  
PI WPI; 2004-652914/63.  
XX  
PT New isolated cancer-associated polynucleotides and polypeptides useful  
PT for diagnosing, preventing or treating cancer, especially lymphoma and  
PT leukemia, or in screening for agents that modulate cancer.  
XX  
PS disclosure; seqid 148; 310pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising at least 10  
CC contiguous nucleotides of any of the 233 polynucleotide sequences given  
CC in the specification, or its complement. The nucleic acids encode cancer-  
CC associated proteins. Also included are an expression vector comprising  
CC the isolated nucleic acid cited above, a host cell comprising the above  
CC recombinant nucleic acid or expression vector, a microarray for detecting  
CC a cancer-associated (CA) nucleic acid comprising at least one probe  
CC comprising at least 10 contiguous nucleotides of any of the above-  
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within  
CC an open reading frame of a CA sequence selected from any of the 95  
CC polynucleotide sequences as mentioned in the specification, or its  
CC complement), an isolated antibody, (or its antigen binding fragment) that  
CC binds to the above polypeptide, a hybridoma that produces the above  
CC monoclonal antibody, a pharmaceutical composition comprising the above  
CC antibody and a pharmaceutical excipient, a kit for detecting cancer  
CC cells/comprising the antibody cited above, methods for diagnosing cancer  
CC or for detecting the presence or absence of cancer cells in an  
CC individual, a method for inhibiting growth of cancer cells in an  
CC individual, a method for delivering a therapeutic agent to cancer cells  
CC in an individual, an electronic library comprising the above  
CC polynucleotide or polypeptide (or their fragments), methods of screening  
CC for anticancer activity or for a bioactive agent capable of modulating  
CC the activity of a CA protein (CAP), methods for detecting cancer  
CC associated with expression of a polypeptide in a test cell sample, a  
CC method for treating cancer and a method for inhibiting the expression of  
CC a CA gene in a cell. The composition and methods are useful for detecting,  
CC diagnosing, preventing and treating cancer, especially lymphoma and  
CC leukemia. These may also be used in screening for agents that modulate  
CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_Pct\_sequences  
XX  
SQ Sequence 421609 BP; 107858A; 90251C; 92730G; 119218T; 0U; 115520Other;

Query Match 83.3%; Score 15; DB 13; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 CCCCCTCTTCTGGAT 16  
DB 55136 CCCCCTCTTCTGGAT 55150  
RESULT 31  
AD209671/c  
ID AD209671 standard; DNA; 224931 BP.  
XX  
AC AD209671;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Human breast cancer marker ZAP3 DNA.  
XX  
KM neoplasm; tumor marker; cytostatic; breast tumor; gene therapy; da; gene;  
XX ZAP3.  
XX  
OS Homo sapiens.  
XX  
PN BP1522594-A2.  
XX  
PD 13-APR-2005.  
XX  
PF 30-JUN-2004; 2004EP-00015374.  
PF  
PR 06-OCT-2003; 2003EP-00022587.  
PR  
XX (FARB ) BAYER HEALTHCARE AG.  
XX  
PI Munnes M, Bojar H;  
PI WPI; 2005-265481/28.  
DR  
DR P-PSDB; AD209836.  
XX  
PT Characterizing the state of a neoplastic disease in a subject by  
PT comparing the pattern of determined expression levels of marker genes in  
PT a biological sample with reference patterns of expression levels.  
XX  
PS Claim 1; SEQ ID NO 115; 542pp; English.  
XX  
CC The invention relates to a novel method for characterizing the state of a  
CC neoplastic disease in a subject comprising determining the pattern of  
CC expression levels of at least 6, 8, 10, 15, 20, 30, 47 or 67 marker genes  
CC in a biological sample, comparing the pattern of expression levels with  
CC reference patterns of expression levels and characterizing the state of  
CC the neoplastic disease in the subject from the outcome of the comparison.  
CC The method of the invention demonstrates cytostatic applications and may  
CC be useful in characterizing the state of a neoplastic disease in a  
CC subject, preferably breast cancer, in gene therapy and for screening for  
CC compounds having a curative effect on a neoplastic disease. The current  
CC sequence is that of the human breast cancer marker ZAP3 DNA of the  
CC invention which is differentially expressed in neoplastic tissue.  
XX  
SQ Sequence 224931 BP; 64720 A; 46885 C; 47309 G; 66017 T; 0 U; 0 Other;  
QY Query Match 83.3%; Score 15; DB 14; Length 224931;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 5403 CCCCCTCTTCTGGAT 5389  
RESULT 32  
ADE53800/c  
ID ADE53800 standard; cDNA; 413 BP.

```
XX AC ADE53800;
XX AC
XX DT 29-JAN-2004 (first entry)
XX XX
XX DE Human prostate cancer CDNA #147.
XX XX
XX KW Human; prostate cancer; s9; CDNA combination; differential expression;
XX gene.
XX OS Homo sapiens.
XX XX
XX PN US2003190640-A1.
XX PD
XX PD 09-OCT-2003.
XX PF
XX PF 29-MAY-2002; 2002US-00252157.
XX PR
XX PR 31-MAY-2001; 2001US-0295048P.
XX PA (PARI/) PARIS M.
XX PA (PEAR/) PEARSON C I.
XX PI
XX PI Farie M, Pearson CI;
XX XX
XX DR WPI; 2003-831619/77.
XX XX
XX PT New combination comprising cDNAs that are differentially expressed in
XX prostate cancer, useful for diagnosing, treating or monitoring the
XX progression of treatment of prostate cancer.
XX XX
XX PS Claim 5; SEQ ID NO 147; 42pp; English.
XX XX
XX CC The invention relates to a combination comprising a number of cDNAs
XX expressed in prostate cancer. The invention also relates to a method for
XX detecting differential expression of one or more cDNAs in a sample
XX containing nucleic acids by hybridising a substrate with the nucleic
XX acids, thus forming one or more hybridisation complexes, detecting
XX hybridisation complex formation and comparing the complexes formed with
XX standard complexes, where differences between the standard and the sample
XX complex formation indicate differential expression of cDNAs in the
XX sample. The differential expression is diagnostic of prostate cancer. The
XX invention also relates to proteins and antibodies related to the cDNAs.
XX CC The combination is useful for diagnosing, treating or monitoring the
XX progression of treatment of prostate cancer. The antibodies are useful
XX for detecting prostate cancer. This sequence represents a human prostate
XX cancer CDNA of the invention.
XX SO
XX SO Sequence 413 BP; 73 A; 128 C; 151 G; 61 T; 0 U; 0 Other;
XX
XX Query Match 82.2%; Score 14.8; DB 10; Length 413;
XX Best Local Similarity 88.9%; Pred. No. 2.1e+03;
XX Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 CCCCCTCTTGGATAC 18
XX |||||||||
XX Db 95 CCCCCTCTGCTGCTAC 78
XX
XX RESULT 33
XX ADQ79063
XX ID ADQ79063 standard; DNA; 616 BP.
XX AC
XX AC ADQ79063;
XX DT
XX DT 21-OCT-2004 (first entry)
XX XX
XX DE Novel canine microarray-related DNA sequence SeqID10622.
XX XX
XX KW canine microarray; drug screening; toxicity assay;
XX environmental pollutant; cellular response; gene expression profile;
XX toxic response; liver necrosis; fatty liver disease;
XX protein adduct formation; hepatitis; dog; ds.
```

```
XX OS Canis familiaris.
XX XX
XX PN WO2004063324-A2.
XX XX
XX PD 29-JUL-2004.
XX PD
XX PF 05-MAY-2003; 2003WO-US013853.
XX XX
XX PR 03-MAY-2002; 2002US-0377240P.
XX XX
XX PA (GENE-) GENE LOGIC INC.
XX PA (PFIZ) PFIZER PROD INC.
XX PI
XX PI Digans JC, Porter M, Wei T;
XX XX
XX DR WPI; 2004-561890/54.
XX XX
XX PT New isolated nucleic acid molecule, useful for drug screening and
XX toxicity assays or for assessing the impact, including toxicity, of a
XX compound, pharmaceutical agent or environmental pollutant on a cell or
XX living organism.
XX PS
XX PS Claim 1; SEQ ID NO 10622; 41pp; English.
XX XX
XX CC This invention is related to a novel isolated canine nucleic acid
XX sequences and the construction of canine microarrays containing a
XX significant portion of the canine genome. The isolated canine nucleic
XX acid sequences of the invention may be useful for drug screening and
XX toxicity assays. The invention is therefore useful for assessing the
XX impact, including toxicity, of a compound, pharmaceutical agent or
XX environmental pollutant on a cell or living organism. The methods are
XX useful for detecting genes that are up- or down-regulated in canines in a
XX disease state. The sequences are useful as diagnostic agents or markers
XX to detect a cellular response in a sample individually or as part of a
XX gene expression profile. It is also useful as a target for agents that
XX modulate gene expression or activity. The database is useful for
XX producing electronic Northern that allow the user to determine the cell
XX type or tissue in which a given gene is expressed and to allow
XX determination of the abundance or expression level of a given gene in a
XX particular tissue or cell. The methods are useful for determining the
XX similarity of a toxic response to one or more individual compounds. The
XX methods are useful for predicting at least one toxic response or the
XX likelihood that a compound or test agent will induce various specific
XX pathologies such as those of the liver (liver necrosis, fatty liver
XX disease, protein adduct formation or hepatitis), those of the kidney,
XX heart, brain or testes, or other pathologies associated with at least one
XX of the toxins. The methods are also useful for predicting or elucidating
XX the potential cellular pathways influenced, induced or modulated by the
XX compound or test agent due to the similarity of the expression profile
XX compared to the profile induced by a known toxin. The present sequence is
XX that of a canine DNA sequence which was claimed for use during the
XX production of a canine microarray of the invention. Note: The sequence
XX data for this patent does not form part of the printed specification but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SO
XX SO Sequence 616 BP; 131 A; 148 C; 166 G; 171 T; 0 U; 0 Other;
XX
XX Query Match 82.2%; Score 14.8; DB 13; Length 616;
XX Best Local Similarity 88.9%; Pred. No. 2.1e+03;
XX Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 CCCCCTCTTGGATAC 18
XX |||||||||
XX Db 48 CCCCCTCTTGGGTAC 65
XX
XX RESULT 34
XX AAS43271/C
XX ID AAS43271 standard; CDNA; 682 BP.
XX AC
XX AC AAS43271;
```



XX 18-DEC-2001 (first entry)  
 DT DNA encoding novel bone marrow polypeptide #6.  
 XX  
 DE  
 XX Human; bone marrow; inflammation; arthritis; nephritis; Crohn's disease;  
 KM ischaemia-reperfusion injury; shock; sepsis; haematopoiesis; bone growth;  
 KM cancer; metastasis; transgenic animal; nerve regeneration; neuropathy;  
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; ALS;  
 KM amyotrophic lateral sclerosis; lymphoid cell disorder; platelet disorder;  
 KM thrombocytopenia; burn; ulcer; osteoporosis; periodontal disease; SCID;  
 KM lung fibrosis; liver fibrosis; immune deficiency; autoimmune disorder;  
 KM severe combined immunodeficiency; infection; multiple sclerosis; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX MO200164840-A2.  
 PN  
 XX 07-SEP-2001.  
 PD  
 XX 28-FEB-2001; 2001MO-US006509.  
 PF  
 XX 28-FEB-2000; 2000US-00515126.  
 PR 18-MAY-2000; 2000US-00577409.  
 PR 30-NOV-2000; 2000US-0250583P.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Ford JE, Boyle BJ, Tang YT, Liu C, Drmanac RT.  
 PT MPI; 2001-582153/65.  
 DR P-PSDB; AAU27328.  
 XX  
 PT Novel bone marrow-expressed nucleic acids and polypeptides for research,  
 PT diagnosis and treatment of hematopoietic, autoimmune, inflammatory  
 PT disorders and cancer and for use in stem cell survival, bone marrow and  
 PT remodeling.  
 XX  
 PS Claim 1; Page 132; 228pp; English.  
 XX  
 XX The invention relates to novel isolated bone marrow-expressed  
 CC polynucleotide (I) (or its complement) comprising a sequence selected  
 CC from 150 sequences of defined base pair sequences given in the  
 CC specification, its mature coding portion or active domain coding portion.  
 CC (I) provided as a collection on a nucleic acid array is useful for  
 CC detecting full-matches or mismatches to any one of the polynucleotides in  
 CC the collection. (I) and its encoded polypeptides (II) are useful for  
 CC treating inflammatory conditions such as arthritis, nephritis, Crohn's  
 CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses,  
 CC and are involved in increasing haematopoiesis, stem cell survival, bone  
 CC growth and remodeling. (II) is involved in cancer cell generation and  
 CC proliferation in metastasis. (I), (II) and modulators of (I) are useful  
 CC for prophylaxis or treatment of cancer. (I) is also useful for creating  
 CC transgenic animals useful for studying the in vivo activities of the  
 CC polypeptide as well as for studying modulators of the polypeptides. (II)  
 CC induces the proliferation of neural cells and regeneration of nerve and  
 CC brain tissue and is useful for the treatment of central and peripheral  
 CC nervous system diseases and neuropathies, such as Alzheimer's,  
 CC Parkinson's disease, Huntington's disease and amyotrophic lateral  
 CC sclerosis. (II) is involved in chemotactic or chemokinetic activity,  
 CC regulation of haematopoiesis and is useful for treating myeloid or  
 CC lymphoid cell disorders, platelet disorders such as thrombocytopenia and  
 CC growth, and in tissue repair, healing of burns, incisions, ulcers, for  
 CC treating osteoporosis, osteoarthritis, bone degenerative disorders or  
 CC periodontal disease. (II) is also useful for gut protection or  
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury  
 CC in various tissues, various immune deficiencies and disorders including  
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
 CC and autoimmune disorders e.g. multiple sclerosis. AA54326-AA54345  
 CC represent human bone marrow polynucleotide sequences of the invention  
 XX  
 S0 Sequence 682 BP; 158 A; 180 C; 192 G; 152 T; 0 U; 0 Other;

Query Match 82.2%; Score 14.8; DB 4; Length 682;  
 Best Local Similarity 88.9%; Pred. No. 2.2e+03;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CCCCCCTTCTGTGATAC 18  
 Db 100 CCCCCCTTCTGTGATAC 83  
 RESULT 35  
 AA543346/C  
 ID AA543346 standard; cDNA; 682 BP.  
 XX  
 AC AA543346;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE DNA encoding novel bone marrow polypeptide #61.  
 XX  
 XX Human; bone marrow; inflammation; arthritis; nephritis; Crohn's disease;  
 KM ischaemia-reperfusion injury; shock; sepsis; haematopoiesis; bone growth;  
 KM cancer; metastasis; transgenic animal; nerve regeneration; neuropathy;  
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; ALS;  
 KM amyotrophic lateral sclerosis; lymphoid cell disorder; platelet disorder;  
 KM thrombocytopenia; burn; ulcer; osteoporosis; periodontal disease; SCID;  
 KM lung fibrosis; liver fibrosis; immune deficiency; autoimmune disorder;  
 KM severe combined immunodeficiency; infection; multiple sclerosis; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX MO200164840-A2.  
 PN  
 XX 07-SEP-2001.  
 PD  
 XX 28-FEB-2001; 2001MO-US006509.  
 PF  
 XX 28-FEB-2000; 2000US-00515126.  
 PR 18-MAY-2000; 2000US-00577409.  
 PR 30-NOV-2000; 2000US-0250583P.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Ford JE, Boyle BJ, Tang YT, Liu C, Drmanac RT.  
 PT MPI; 2001-582153/65.  
 DR P-PSDB; AAU27403.  
 XX  
 PT Novel bone marrow-expressed nucleic acids and polypeptides for research,  
 PT diagnosis and treatment of hematopoietic, autoimmune, inflammatory  
 PT disorders and cancer and for use in stem cell survival, bone marrow and  
 PT remodeling.  
 XX  
 PS Claim 1; Page 181; 228pp; English.  
 XX  
 XX The invention relates to novel isolated bone marrow-expressed  
 CC polynucleotide (I) (or its complement) comprising a sequence selected  
 CC from 150 sequences of defined base pair sequences given in the  
 CC specification, its mature coding portion or active domain coding portion.  
 CC (I) provided as a collection on a nucleic acid array is useful for  
 CC detecting full-matches or mismatches to any one of the polynucleotides in  
 CC the collection. (I) and its encoded polypeptides (II) are useful for  
 CC treating inflammatory conditions such as arthritis, nephritis, Crohn's  
 CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses,  
 CC and are involved in increasing haematopoiesis, stem cell generation and  
 CC growth and remodeling. (II) is involved in cancer cell generation and  
 CC proliferation in metastasis. (I), (II) and modulators of (I) are useful  
 CC for prophylaxis or treatment of cancer. (I) is also useful for creating  
 CC transgenic animals useful for studying the in vivo activities of the  
 CC polypeptide as well as for studying modulators of the polypeptides. (II)  
 CC induces the proliferation of neural cells and regeneration of nerve and  
 CC brain tissue and is useful for the treatment of central and peripheral  
 CC nervous system diseases and neuropathies, such as Alzheimer's,

CC parkinson's disease, Huntington's disease and amyotrophic lateral  
CC sclerosis. (ii) is involved in chemotactic or chemokinetic activity,  
CC regulation of haematopoiesis and is useful for treating myeloid or  
CC lymphoid cell disorders, platelet disorders such as thrombocytopenia and  
CC for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue  
CC growth, and in tissue repair, healing of burns, incisions, ulcers, for  
CC treating osteoporosis, osteoarthritis, bone degenerative disorders or  
CC periodontal disease. (iii) is also useful for gut protection or  
CC regeneration and treatment of lung or liver fibrosis, reperfusion injury  
CC in various tissues, various immune deficiencies and disorders including  
CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
CC and autoimmune disorders e.g. multiple sclerosis. AAS43266-AAS43415  
CC represent human bone marrow polynucleotide sequences of the invention  
XX

SQ Sequence 682 BP; 157 A; 180 C; 192 G; 152 T; 0 U; 1 Other;

Query Match 82.2%; Score 14.8; DB 4; Length 682;  
Best Local Similarity 88.9%; Pred. No. 2.2e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCCCCTCTTGATAC 18  
|||  
100 CCCCCCTCTTGATAC 83

Db

RESULT 36  
AA185744/c  
ID AA185744 standard; cDNA; 682 BP.  
XX  
AC AA185744;  
XX  
XX 06-NOV-2001 (first entry)  
XX  
XX Human polynucleotide SEQ ID NO 5804.  
DE  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200164835-A2.  
PN  
XX 07-SEP-2001.  
PD  
XX 26-FEB-2001; 2001WO-US004927.  
PF  
XX 28-FEB-2000; 2000US-00515126.  
PR  
XX 18-MAY-2000; 2000US-00577409.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-514838/56.  
DR  
XX P-PSDB; AAO05813.  
DR  
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.  
XX  
XX Claim 1; SEQ ID NO 5804; 1399pp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to human polynucleotides (AA179941-AA193841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX

SQ Sequence 682 BP; 157 A; 180 C; 192 G; 152 T; 0 U; 1 Other;

Query Match 82.2%; Score 14.8; DB 4; Length 682;  
Best Local Similarity 88.9%; Pred. No. 2.2e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCCCCTCTTGATAC 18  
|||  
100 CCCCCCTCTTGATAC 83

Db

RESULT 37  
ABL26119/c  
ID ABL26119 standard; DNA; 1095 BP.  
XX  
XX ABL26119;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 29830.  
DE  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX 27-SEP-2001.  
PD  
XX 23-MAR-2001; 2001WO-US009231.  
PF  
XX 23-MAR-2000; 2000US-0191637P.  
PR  
XX 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEXE ) PE CORP NY.  
PA  
XX Venter JC, Adams M, Li PMD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
DR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions.  
XX  
XX Claim 1; SEQ ID NO 29830; 21pp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABR01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABR22072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1095 BP; 343 A; 214 C; 265 G; 273 T; 0 U; 0 Other;

Query Match 82.2%; Score 14.8; DB 4; Length 1095;  
Best Local Similarity 88.9%; Pred. No. 2.3e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCCCCTCTTGATAC 18  
|||  
188 CCCCCCTCTTGATAC 171

Db



PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140035P.  
PR 23-JUN-1999; 99US-0140035P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145513P.  
PR 27-JUL-1999; 99US-0145518P.  
PR 27-JUL-1999; 99US-0145519P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147933P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148655P.  
PR 13-AUG-1999; 99US-0148686P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.

PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-015138P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.2%; Score 14.8; DB 3; Length 1408;  
Best Local Similarity 88.9%; Pred. No. 2.3e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGATAC 18  
||| |||||  
Db 931 CCCATCTTCTGATAC 914

RESULT 40  
ADT19869/c  
ID ADT19869 standard; cDNA; 2297 BP.  
XX  
AC ADT19869;  
XX  
DT 13-JAN-2005 (first entry)  
XX  
DE Plant cDNA, Seq ID 5195.

XX plant; ss; gene; transgenic; cold tolerance; growth rate;  
 KW drought tolerance; disease resistance; galactomannan production;  
 KW plant growth regulator; heat tolerance; herbicide tolerance;  
 KW lignin production; extreme osmotic condition tolerance;  
 KW pathogens resistance; pest resistance; yield improvement; seed oil yield;  
 KW seed protein yield.  
 XX Viridiplantae.  
 XX US2004216190-A1.  
 XX 28-OCT-2004.  
 XX 18-DEC-2003; 2003US-00739930.  
 XX 28-APR-2003; 2003US-00424599.  
 XX 28-APR-2003; 2003US-00425115.  
 XX (KOVA/) KOVALIC D. K.  
 XX Kovalic DK;  
 XX WPI; 2004-757369/74.  
 XX New recombinant DNA constructs useful in the field of biochemistry and  
 PT genetic, and in particular for producing transgenic plants with improved  
 PT biological characteristics.  
 XX Claim 1; SEQ ID NO 5195; 14pp; English.  
 XX The invention relates a recombinant DNA construct comprising a  
 CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:  
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences  
 CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,  
 CC Arabidopsis, wheat and rape but the specification does not indicate which  
 CC sequences is derived from which organism. Also included is a method of  
 CC producing a plant having an improved property, comprising transforming a  
 CC plant with a recombinant DNA construct comprising a promoter region  
 CC functional in a plant cell operably joined to a polynucleotide encoding a  
 CC polypeptide associated with the property, and growing the transformed  
 CC plant. The property is selected from improving plant cold tolerance, for  
 CC manipulating growth rate in plant cells by modification of the cell cycle  
 CC pathway, for improving plant drought tolerance, for providing increased  
 CC resistance to plant disease, for galactomannan production, for production  
 CC of plant growth regulators, for improving plant heat tolerance, for  
 CC improving plant tolerance to herbicides, for increasing the rate of  
 CC homologous recombination in plants, for lignin production, for improving  
 CC plant tolerance to extreme osmotic conditions, for improving plant  
 CC tolerance to pathogens or pests, for yield improvement by modification of  
 CC photosynthesis, for modifying seed oil yield and/or content, for  
 CC modifying seed protein yield and/or content, for yield improvement by  
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
 CC and for yield improvement by providing improved plant growth and  
 CC development under at least one stress condition. The polynucleotide may  
 CC also encode a plant transcription factor. The methods and compositions of  
 CC the present invention are useful in the field of biochemistry and  
 CC genetics, in particular for producing transgenic plants with improved  
 CC biological characteristics such as increased yield, improved nitrogen  
 CC flow, increasing plant tolerance to cold or heat, improving plant  
 CC tolerance to extreme osmotic and drought conditions, and improving plant  
 CC tolerance to plant pests or pathogens. They can also be used in physical  
 CC arrays of molecules, plant breeding markers, computer-based storage and  
 CC analysis systems. The present sequence is one of the 5544 plant cDNA  
 CC sequences of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docID=20040216190.  
 XX Sequence 2297 BP; 628 A; 481 C; 577 G; 611 T; 0 U; 0 Other;

Query Match 82.2%; Score 14.8; DB 13; Length 2297;  
 Best Local Similarity 88.9%; Pred. No. 2.4e+03;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CCCCCCTCTTCGATAC 18  
 DB 2150 CACCCCTCTTCGATAC 2133  
 RESULT 41  
 AAS79005/c  
 ID AAS79005 standard; cDNA; 2433 BP.  
 XX AAS79005;  
 XX 13-FEB-2002 (first entry)  
 XX DNA encoding novel human diagnostic protein #14809.  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX Homo sapiens.  
 XX WO200175067-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US008631.  
 XX 31-MAR-2000; 2000US-00540217.  
 XX 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Dmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 XX P-PSDB; ABG14818.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 14809; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 2433 BP; 545 A; 683 C; 741 G; 464 T; 0 U; 0 Other;

Query Match 82.2%; Score 14.8; DB 5; Length 2433;  
 Best Local Similarity 88.9%; Pred. No. 2.4e+03;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTTGATAC 18  
 |||||  
 Db 168 CCCACCTCTTCTGACAC 151

RESULT 42

ABL26118/c  
 ID ABL26118 standard; DNA; 3095 BP.

AC ABL26118;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 29827.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

XX interactions.

PS Claim 1; SEQ ID NO 29827; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signaling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-

CC ABBS72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

CC Sequence 3095 BP; 962 A; 638 C; 659 G; 836 T; 0 U; 0 Other;

XX Query Match 82.2%; Score 14.8; DB 4; Length 3095;

XX Best Local Similarity 88.9%; Pred. No. 2.5e+03;

XX Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTTGATAC 18  
 |||||  
 Db 1188 CCCGCTCTTCTGATCC 1171

RESULT 43

AD063840  
 ID AD063840 standard; cDNA; 3788 BP.

AC AD063840;

DT 07-OCT-2004 (first entry)

XX Novel human cDNA sequence #1001.

KW ss; gene; osteopathic; neuroprotective; noctropic; antiparkinsonian;  
 KW cytoskeletal; gene therapy; diagnostic marker; morbid state; osteoporosis;  
 KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;  
 KW cancer.

XX Homo sapiens.

XX EPI440981-A2.

XX 28-JUL-2004.

XX 21-JAN-2004; 2004EP-00001196.

XX 21-JAN-2003; 2003JP-00102206.

XX 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otaki T, Wakamatsu A, Sato H, Ishii S;

XX Yamamoto J, Isono Y, Nagai K, Irie R;

XX WPI; 2004-535376/52.

XX P-PSDB; AD066028.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,

XX Alzheimer's disease, Parkinson's disease, dementia and various cancers.

XX Claim 1; SEQ ID NO 1001; 2449pp; English.

XX The invention relates to 2495 novel polynucleotides (1) and their encoded

CC polypeptides, sequences hybridizing to these nucleotides, sequences

CC encoding partial polypeptide and sequences having 70% or 90% identity to

CC the nucleotide and protein sequences. The nucleotides and polypeptides

CC are useful as diagnostic markers or therapeutic target for the diseases

CC or morbid states. They are also useful for treating osteoporosis,

CC neurological diseases, Alzheimer's disease, Parkinson's disease,

CC dementia and various cancers. This sequence corresponds to a nucleotide

CC sequence of the invention.

XX Sequence 3788 BP; 738 A; 1093 C; 1095 G; 862 T; 0 U; 0 Other;

XX Query Match 82.2%; Score 14.8; DB 12; Length 3788;

XX Best Local Similarity 88.9%; Pred. No. 2.5e+03;

XX Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTTGATAC 18  
 |||||  
 Db 328 CCCGCTCTGCTGCTAC 345

RESULT 44

AAK73805/c  
 ID AAK73805 standard; DNA; 3861 BP.

AC AAK73805;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28617.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytoskeletal; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249266P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
  
(HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX  
PS Disclosure; SEQ ID NO 26617; 3071bp + Sequence Listing; English.  
XX

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins, and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
CC  
SQ Sequence 3861 BP; 840 A; 1033 C; 1048 G; 940 T; 0 U; 0 Other;  
  
Query Match 82.2%; Score 14.8; DB 4; Length 3861;  
Best Local Similarity 88.9%; Pred. No. 2.5e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 CCCCCTCTTGATAC 18  
Db 1891 CCCACCTCTTGACAC 1874  
  
RESULT 45  
AAK73521/C  
ID AAK73521 standard; DNA; 3861 BP.  
XX AAK73521;  
AC AAK73521;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28333.  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
FN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225471P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233633P.  
PR 14-SEP-2000; 2000US-0233664P.  
PR 14-SEP-2000; 2000US-0233665P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.





```
XX XX Retroviral vector; inflammatory disorder; dermatological disorder;
KM cardiovascular disorder; autoimmune disease; neurological disorder;
KM cancer; gene therapy; ss.
XX OS Synthetic.
XX PN WO200056910-A1.
XX PD 28-SEP-2000.
XX PF 22-MAR-2000; 2000WO-GB001091.
XX PR 22-MAR-1999; 99GB-00006615.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Uden M, Mltrophanus K;
XX DR WPI; 2000-628271/60.
XX PT Retroviral vector for delivering one or more nucleotides of interest to a
PT target site has functional and non-functional splice donor and splice
PT acceptor sites.
XX PS Example 8, Fig 33, 148pp; English.
XX CC The specification describes a retroviral vector which comprises a
XX CC nucleotide sequence of interest flanked by functional splice donor sites
XX CC and functional splice acceptor sites. The vector is derived from a
XX CC retroviral pro-vector. The retroviral vector is useful for preparing
XX CC pharmaceutical compositions to deliver one or more nucleotide sequences
XX CC of interest to a target site. The retroviral vectors are especially useful
XX CC for treating inflammatory disorders, cancers, dermatological disorders,
XX CC cardiovascular disorders, autoimmune diseases and neurological disorders.
XX CC The retroviral vector is useful in gene therapy. The present sequence
XX CC represents the retroviral vector pTRONIN, a vector of the invention
XX SQ Sequence 9830 BP; 2209 A; 2659 C; 2688 G; 2274 T; 0 U; 0 Other;
Query Match 82.2%; Score 14.8; DB 3; Length 9830;
Best Local Similarity 88.9%; Pred. No. 2.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCCCCCTCTTCGATAC 18
DB 6112 CCCCCCTTTCTGAGAC 6095
RESULT 48
AAV49655/c
ID AAV49655 standard; DNA; 28720 BP.
XX AC AAV49655;
XX DT 23-OCT-1998 (first entry)
XX DE Human SC3 DNA.
XX DE Human SC3 DNA.
XX KM Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;
KM nervous system; medullo-blastoma; glioma; breast; detection; SC3;
KM autantbody; ss.
XX OS Homo sapiens.
XX PN WO9830687-A2.
XX PD 16-JUL-1998.
XX PF 09-JAN-1998; 98WO-DE000096.
XX PR 09-JAN-1997; 97DE-01000519.
XX PR 18-JUL-1997; 97DE-01030997.
```

```
XX XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PI Mollenhauer J, Pouetka A;
XX DR WPI; 1998-399136/34.
XX PT Proteins containing scavenger receptor, cysteine rich domain - useful for
PT diagnosis and treatment of tumors.
XX PS Disclosure; Fig 5c; 54pp; German.
XX CC This nucleotide sequence contains a fragment of a gene which is capable
XX CC of encoding a protein which contains a SRCR (scavenger receptor, cysteine
XX CC rich) domain. The gene and encoded protein can be used to diagnose or
XX CC treat tumors, particularly of the nervous system (medullo-blastoma or
XX CC glioma) or breast. The DNA sequence and probes derived from it, are used
XX CC to identify genes that express SRCR-domain containing proteins, to
XX CC determine the form in which these proteins exist and to assess the
XX CC significance of individual forms on cellular properties. The protein can
XX CC be used to detect the presence of autoantibodies and antibodies which
XX CC regulate its expression
XX SQ Sequence 28720 BP; 6956 A; 6728 C; 6982 G; 8054 T; 0 U; 0 Other;
Query Match 82.2%; Score 14.8; DB 2; Length 28720;
Best Local Similarity 88.9%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCCCCCTCTTCGATAC 18
DB 1691 CCTGCCTCTTCGATAC 1674
RESULT 49
ACN37225
ID ACN37225 standard; DNA; 30135 BP.
XX AC ACN37225;
XX DT 18-NOV-2004 (first entry)
XX DE Human periodontal disease related gene COL1A1 SEQ ID NO:135.
XX KM periodontal disease; polymorphism; ds; human; gene; SNP;
XX KM single nucleotide polymorphism.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT misc_feature 109
XX FT /tag= a
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT /note= "Variable nucleotide C,T"
XX FT 654
XX FT /tag= b
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT /note= "Variable nucleotide G,A"
XX FT 732
XX FT /tag= c
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT /note= "Variable nucleotide G,A"
XX FT 1798
XX FT /tag= d
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT /note= "Variable nucleotide C,G"
XX FT 2708
XX FT /tag= e
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT /note= "Variable nucleotide G,A"
XX FT 2903
XX FT /tag= f
XX FT /standard_name= "Single nucleotide polymorphism"
```

```
PT /note= "Variable nucleotide G,T"
PT misc_feature
PT 4079
PT /*tag= g
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,G"
PT
PT misc_feature
PT 4228
PT /*tag= h
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,T"
PT
PT misc_feature
PT 4252
PT /*tag= i
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,C"
PT
PT misc_feature
PT 4792
PT /*tag= j
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide C,T"
PT
PT misc_feature
PT 5120
PT /*tag= k
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide C,T"
PT
PT misc_feature
PT 5301
PT /*tag= l
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide C,T"
PT
PT misc_feature
PT 7409
PT /*tag= m
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide C,G"
PT
PT misc_feature
PT 7610
PT /*tag= n
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,G"
PT
PT misc_feature
PT 8346
PT /*tag= o
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,G"
PT
PT misc_feature
PT 8373
PT /*tag= p
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,G"
PT
PT misc_feature
PT 8690
PT /*tag= q
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide G,A"
PT
PT misc_feature
PT 8832
PT /*tag= r
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide G,A"
PT
PT misc_feature
PT 9099
PT /*tag= s
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide G,A"
PT
PT misc_feature
PT 9755
PT /*tag= t
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,T"
PT
PT misc_feature
PT 9786
PT /*tag= u
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,C"
PT
PT misc_feature
PT 9829
PT /*tag= v
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide T,G"
PT
PT misc_feature
PT 9942
PT /*tag= w
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide T,G"
PT
PT misc_feature
PT 9947
PT /*tag= x
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide T,G"
PT
PT misc_feature
PT 10001
PT /*tag= y
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide C,T"
PT
PT misc_feature
PT 10069
PT /*tag= z
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,T"
PT
PT misc_feature
PT 10196
PT /*tag= aa
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide T,G"
PT
PT misc_feature
PT 10232
PT /*tag= ab
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide T,G"
PT
PT misc_feature
PT 10519
PT /*tag= ac
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide T,G"
PT
PT misc_feature
PT 10883
PT /*tag= ad
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,G"
PT
PT misc_feature
PT 10977
PT /*tag= ae
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,T"
PT
PT misc_feature
PT 11043
PT /*tag= af
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide C,G"
PT
PT misc_feature
PT 11045
PT /*tag= ag
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide T,G"
PT
PT misc_feature
PT 11310
PT /*tag= ah
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide T,G"
PT
PT misc_feature
PT 11622
PT /*tag= ai
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,G"
PT
PT misc_feature
PT 11697
PT /*tag= aj
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide C,T"
PT
PT misc_feature
PT 12189
PT /*tag= ak
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,G"
PT
PT misc_feature
PT 12406
PT /*tag= al
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide A,G"
PT
PT misc_feature
PT 12475
PT /*tag= am
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide T,C"
PT
PT misc_feature
PT 12704
PT /*tag= an
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide G,A"
PT
PT misc_feature
PT 13649
PT /*tag= ao
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide C,G"
PT
PT misc_feature
PT 14271
PT /*tag= ap
PT /standard_name= "Single nucleotide polymorphism"
PT /note= "Variable nucleotide C,G"
PT
PT misc_feature
PT 14434
```

```
FT FT /tag= ag
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT /note= "Variable nucleotide C,G"
FT FT misc_feature
FT FT /tag= ar
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT /note= "Variable nucleotide A,C"
FT FT misc_feature
FT FT /tag= as
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT /note= "Variable nucleotide G,A"
FT FT misc_feature
FT FT /tag= at
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT /note= "Variable nucleotide T,G"
FT FT misc_feature
FT FT /tag= au

Query Match 82.2%; Score 14.8; DB 13; Length 30135;
Best Local Similarity 88.9%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATAC 18
Db 22714 CCCACTCTCTGATAC 22731

RESULT 50
ADQ97934
ID ADQ97934 standard; DNA; 69648 BP.
AC ADQ97934;
DT 07-OCT-2004 (first entry)
XX Human cancer associated sequence HD11-025, SEQ ID 911.
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX Homo sapiens.
XX WO2004060304-A2.
XX 22-JUL-2004.
XX 22-DEC-2003; 2003WO-US041389.
XX PF 27-DEC-2002; 2002US-00330773.
XX PR (SAGR-) SAGRES DISCOVERY INC.
XX PA (SAGR-) SAGRES DISCOVERY INC.
XX PI Morris DW, Malandro MS;
XX WPI; 2004-543781/52.
XX 2004-543781/52.
XX New isolated cancer associated nucleic acids comprising at least 10
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX cancers such as leukemia and lymphoma.
XX Claim 1; SEQ ID NO 911; 199pp; English.
XX The present invention relates to cancer associated sequences (ADQ97025-
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 69648 BP; 20376 A; 13878 C; 15258 G; 20136 T; 0 U; 0 Other;

Query Match 82.2%; Score 14.8; DB 12; Length 69648;
Best Local Similarity 88.9%; Pred. No. 3.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CCCCCCTCTTGATAC 18
Db 21264 CCCACTCTCTGATAC 21281

RESULT 51
ID ACN45192 standard; DNA; 82660 BP.
XX ACN45192;
XX ACN45192;
XX 18-NOV-2004 (first entry)
XX Mouse genomic sequence MCG9397.
XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX Mus musculus.
XX WO2003073826-A2.
XX 12-SEP-2003.
XX 28-FEB-2003; 2003WO-US006235.
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
XX Claim 1; SEQ ID NO 2017; opp; English.
XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX carcinoma; (vi) for inhibiting the activity of CAP; (ix) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published
XX Sequence 82660 BP; 19374 A; 15352 C; 16570 G; 20443 T; 0 U; 10921 Other;

Query Match 82.2%; Score 14.8; DB 11; Length 82660;
Best Local Similarity 88.9%; Pred. No. 3.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATAC 18
Db 57791 CCCACTCTCTGATAC 57808

RESULT 52
ID AD213310 standard; DNA; 83493 BP.
XX AD213310;
XX AD213310;
XX 16-JUN-2005 (first entry)
```

DE Murine cancer-associated genomic DNA #71.  
XX  
XX Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;  
KM cytostatic; gene; ds.  
XX  
XX Mus sp.  
XX WO2005031001-A2.  
XX  
XX 07-APR-2005.  
XX  
XX 23-SEP-2004; 2004WO-US031617.  
XX  
XX 23-SEP-2003; 2003US-00669920.  
XX  
XX (CHIR ) CHIRON CORP.  
XX  
XX Morris DW, Malandro MS;  
XX WPI; 2005-273395/28.  
XX  
XX Nucleic acid array useful for detecting cancer associated nucleic acid,  
PT comprises two or more nucleic acid probes.  
XX  
XX  
XX Disclosure; SEQ ID NO 830; 198pp; English.  
XX  
XX The invention relates to a nucleic acid array for detecting a cancer  
CC associated (CA) nucleic acid, comprising two or more nucleic acid probes.  
CC The invention also relates to a peptide array comprising two or more  
CC isolated polypeptides encoded by a CA nucleic acid sequence, a compound  
CC that binds to a polypeptide, an isolated antibody or its fragment which  
CC binds to a polypeptide, which is prepared by immunizing a host animal  
CC with a composition comprising the polypeptide or its antigen binding  
CC fragment and collecting cells from the host expressing antibodies against  
CC the antigen or its antigen binding fragment, a composition comprising the  
CC antibody and a carrier, a method of screening for anticancer activity, a  
CC method of detecting a CA nucleic acid, a method of diagnosing cancer, a  
CC method of treating cancer and a method of inhibiting expression of a CA  
CC nucleic acid in a cell. The CA nucleic acids are useful for detecting CA  
CC nucleic acid. The antibody is useful for detecting the presence or  
CC absence of cancer cells in an individual which involves contacting cells  
CC from the individual with the antibody and detecting a complex of a CA  
CC protein from the cancer cells and the antibody, where the detection of  
CC the complex correlates with the presence of cancer cells in the  
CC individual. The composition is useful for inhibiting growth of cancer  
CC cells in an individual or for delivering a therapeutic agent to cancer  
CC cells in an individual. The invention is also useful for diagnosing  
CC cancer, for treating cancer and for inhibiting expression of a CA gene in  
CC a cell. This sequence represents murine cancer-associated genomic DNA of  
CC the invention.  
XX  
SQ Sequence 83493 BP; 22528 A; 17868 C; 19019 G; 24033 T; 0 U; 45 Other;  
Query Match 82.2%; Score 14.8; DB 14; Length 83493;  
Best Local Similarity 88.9%; Pred. No. 3.1e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CCCCCCTTCTTGATAC 18  
Db 58649 CCCCCCTTCTTGATAC 58666  
RESULT 53  
ADQ97718  
ID ADQ97718 standard; DNA; 109147 BP.  
XX  
XX ADQ97718;  
XX  
XX 07-OCT-2004 (first entry)  
XX  
XX Mouse cancer associated sequence MD10-033, SEQ ID 695.  
XX  
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.

XX  
XX Mus musculus.  
XX  
XX WO2004060304-A2.  
XX  
XX  
XX 22-JUL-2004.  
XX  
XX  
XX 22-DEC-2003; 2003WO-US041389.  
XX  
XX 27-DEC-2002; 2002US-00330773.  
XX  
XX (SAGR-) SAGRES DISCOVERY INC.  
XX  
XX Morris DW, Malandro MS;  
XX WPI; 2004-543781/52.  
XX  
XX  
XX New isolated cancer associated nucleic acids comprising at least 10  
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating  
PT cancers such as leukemia and lymphoma.  
XX  
XX  
XX Claim 1; SEQ ID NO 695; 199pp; English.  
XX  
XX The present invention relates to cancer associated sequences (ADQ97025-  
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or  
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 109147 BP; 26811 A; 25605 C; 26195 G; 28842 T; 0 U; 1694 Other;  
Query Match 82.2%; Score 14.8; DB 12; Length 109147;  
Best Local Similarity 88.9%; Pred. No. 3.2e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CCCCCCTTCTTGATAC 18  
Db 11023 CCCCCCTTCTTGATAC 11040  
RESULT 54  
ADN46845\_18  
Continuation (19 of 21) of ADN46845 from base 1800001 (Thermococcus kodakaraensis KOD1  
WP Sequence Split into 21 fragments LOCUS ADN46845 Accession Adn46845  
WP Fragment Name Begin End  
WP ADN46845\_00 1 110000  
WP ADN46845\_01 100001 210000  
WP ADN46845\_02 200001 310000  
WP ADN46845\_03 300001 410000  
WP ADN46845\_04 400001 510000  
WP ADN46845\_05 500001 610000  
WP ADN46845\_06 600001 710000  
WP ADN46845\_07 700001 810000  
WP ADN46845\_08 800001 910000  
WP ADN46845\_09 900001 1010000  
WP ADN46845\_10 1000001 1110000  
WP ADN46845\_11 1100001 1210000  
WP ADN46845\_12 1200001 1310000  
WP ADN46845\_13 1300001 1410000  
WP ADN46845\_14 1400001 1510000  
WP ADN46845\_15 1500001 1610000  
WP ADN46845\_16 1600001 1710000  
WP ADN46845\_17 1700001 1810000  
WP ADN46845\_18 1800001 1910000  
WP ADN46845\_19 1900001 2010000  
WP ADN46845\_20 2000001 2089378  
Query Match 82.2%; Score 14.8; DB 12; Length 110000;  
Best Local Similarity 88.9%; Pred. No. 3.2e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CCCCCCTTCTTGATAC 18

Db 47484 CTCCTCTTCGTGATTC 47501

## RESULT 55

ADN47591\_02/c

Continuation (13 of 21) of ADN47591 from base 200001 (Thermococcus kodakaraensis KOD1 gen  
WP Sequence split into 21 fragments LOCUS ADN47591 Accession Adn47591

WP	Fragment Name	Begin	End
WP	ADN47591_00	1	110000
WP	ADN47591_01	100001	210000
WP	ADN47591_02	200001	310000
WP	ADN47591_03	300001	410000
WP	ADN47591_04	400001	510000
WP	ADN47591_05	500001	610000
WP	ADN47591_06	600001	710000
WP	ADN47591_07	700001	810000
WP	ADN47591_08	800001	910000
WP	ADN47591_09	900001	1010000
WP	ADN47591_10	1000001	1110000
WP	ADN47591_11	1100001	1210000
WP	ADN47591_12	1200001	1310000
WP	ADN47591_13	1300001	1410000
WP	ADN47591_14	1400001	1510000
WP	ADN47591_15	1500001	1610000
WP	ADN47591_16	1600001	1710000
WP	ADN47591_17	1700001	1810000
WP	ADN47591_18	1800001	1910000
WP	ADN47591_19	1900001	2010000
WP	ADN47591_20	2000001	2089378

Query Match Best Local Similarity 82.2%; Score 14.8; DB 12; Length 110000;  
Pred. No. 3.2e+03; Mismatches 2; Indels 0; Gaps 0;

Matches 16; Conservative 0; Db 41894 CTCCTCTTCGTGATTC 41877

## RESULT 56

ADN46123\_18

Continuation (19 of 21) of ADN46123 from base 1800001 (Thermococcus kodakaraensis KOD1 gen  
WP Sequence split into 21 fragments LOCUS ADN46123 Accession Adn46123

WP	Fragment Name	Begin	End
WP	ADN46123_00	1	110000
WP	ADN46123_01	100001	210000
WP	ADN46123_02	200001	310000
WP	ADN46123_03	300001	410000
WP	ADN46123_04	400001	510000
WP	ADN46123_05	500001	610000
WP	ADN46123_06	600001	710000
WP	ADN46123_07	700001	810000
WP	ADN46123_08	800001	910000
WP	ADN46123_09	900001	1010000
WP	ADN46123_10	1000001	1110000
WP	ADN46123_11	1100001	1210000
WP	ADN46123_12	1200001	1310000
WP	ADN46123_13	1300001	1410000
WP	ADN46123_14	1400001	1510000
WP	ADN46123_15	1500001	1610000
WP	ADN46123_16	1600001	1710000
WP	ADN46123_17	1700001	1810000
WP	ADN46123_18	1800001	1910000
WP	ADN46123_19	1900001	2010000
WP	ADN46123_20	2000001	2089378

Query Match Best Local Similarity 82.2%; Score 14.8; DB 12; Length 110000;  
Pred. No. 3.2e+03; Mismatches 2; Indels 0; Gaps 0;

Matches 16; Conservative 0; Db 41894 CTCCTCTTCGTGATTC 41877

Db 47484 CTCCTCTTCGTGATTC 47501

## RESULT 57

ADN47209\_02/c

Continuation (13 of 21) of ADN47209 from base 200001 (Thermococcus kodakaraensis KOD1 gen  
WP Sequence split into 21 fragments LOCUS ADN47209 Accession Adn47209

WP	Fragment Name	Begin	End
WP	ADN47209_00	1	110000
WP	ADN47209_01	100001	210000
WP	ADN47209_02	200001	310000
WP	ADN47209_03	300001	410000
WP	ADN47209_04	400001	510000
WP	ADN47209_05	500001	610000
WP	ADN47209_06	600001	710000
WP	ADN47209_07	700001	810000
WP	ADN47209_08	800001	910000
WP	ADN47209_09	900001	1010000
WP	ADN47209_10	1000001	1110000
WP	ADN47209_11	1100001	1210000
WP	ADN47209_12	1200001	1310000
WP	ADN47209_13	1300001	1410000
WP	ADN47209_14	1400001	1510000
WP	ADN47209_15	1500001	1610000
WP	ADN47209_16	1600001	1710000
WP	ADN47209_17	1700001	1810000
WP	ADN47209_18	1800001	1910000
WP	ADN47209_19	1900001	2010000
WP	ADN47209_20	2000001	2089378

Query Match Best Local Similarity 82.2%; Score 14.8; DB 12; Length 110000;  
Pred. No. 3.2e+03; Mismatches 2; Indels 0; Gaps 0;

Matches 16; Conservative 0; Db 41894 CTCCTCTTCGTGATTC 41877

## RESULT 58

ADN46464\_18

Continuation (19 of 21) of ADN46464 from base 1800001 (Thermococcus kodakaraensis KOD1 gen  
WP Sequence split into 21 fragments LOCUS ADN46464 Accession Adn46464

WP	Fragment Name	Begin	End
WP	ADN46464_00	1	110000
WP	ADN46464_01	100001	210000
WP	ADN46464_02	200001	310000
WP	ADN46464_03	300001	410000
WP	ADN46464_04	400001	510000
WP	ADN46464_05	500001	610000
WP	ADN46464_06	600001	710000
WP	ADN46464_07	700001	810000
WP	ADN46464_08	800001	910000
WP	ADN46464_09	900001	1010000
WP	ADN46464_10	1000001	1110000
WP	ADN46464_11	1100001	1210000
WP	ADN46464_12	1200001	1310000
WP	ADN46464_13	1300001	1410000
WP	ADN46464_14	1400001	1510000
WP	ADN46464_15	1500001	1610000
WP	ADN46464_16	1600001	1710000
WP	ADN46464_17	1700001	1810000
WP	ADN46464_18	1800001	1910000
WP	ADN46464_19	1900001	2010000
WP	ADN46464_20	2000001	2089378

Query Match Best Local Similarity 82.2%; Score 14.8; DB 12; Length 110000;  
Pred. No. 3.2e+03; Mismatches 2; Indels 0; Gaps 0;

Matches 16; Conservative 0; Db 47484 CTCCTCTTCGTGATTC 47501







```
XX
DT 24-OCT-2001 (first entry)
XX
XX Human surfactant protein B, SPB, TTF-1 containing oligonucleotide Bcm.
DE Human surfactant protein B; SPB; promoter; nuclear protein binding site;
XX Thyroid transcription factor; TTF-1; lung cancer; thyroid cancer; ds;
KW Bcm; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH replace(10. .13, CAAg)
FT mutation /*tag= a
FT /*note= "Mutates the TTF-1 binding site"
XX
XX US2001016352-A1.
XX
XX 23-AUG-2001.
XX
XX 26-MAY-1999; 99US-00320337.
XX
XX 18-MAY-1994; 94US-00245356.
XX 17-MAY-1995; 95US-00442809.
XX
XX (BOHI/) BOHINSKI R J.
XX (WHIT/) WHITSETT J A.
XX
XX Bohinski RJ, Whitsett JA;
XX
XX WPI; 2001-513959/56.
XX
XX Oligonucleotide sequences which bind nuclear proteins and surfactants
PT found in lung cells, useful for detecting cancers that originate in the
PT lung.
XX
XX Example 5; Fig 34a; 76pp; English.
XX
XX The invention relates to an oligonucleotide which includes at least 1
CC nucleic acid sequence which binds to at least 1 nuclear protein found in
CC lung cells (e.g. the thyroid transcription factor 1, TTF-1, protein). The
CC oligonucleotide can be expressed in lung cells via a vector and can be
CC used to target therapeutic agents to kill lung or thyroid cancer cells.
CC The oligonucleotide can be used to detect or diagnose lung or thyroid
CC cancer. The oligonucleotides may be designed from the sequences of, for
CC example, the promoters of lung-specific genes such as those encoding
CC surfactant proteins. The present sequence is a mutated oligonucleotide
CC from human surfactant protein B, SPB, promoter used in an EMSA
CC (electrophoretic mobility shift assay) assay to confirm the presence of
CC TTF-1 binding sites in the SPB promoter. The TTF-1 site has been mutated
XX
SQ Sequence 30 BP; 6 A; 6 C; 12 G; 6 T; 0 U; 0 Other;
Query Match 80.0%; Score 14.4; DB 4; Length 30;
Best Local Similarity 93.8%; Pred. No. 2.6e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 CCCCCTCTTGTGATA 17
Db 24 CCCCCTCTACTGTGATA 9
RESULT 64
AAL15054
ID AAL15054 standard; cDNA; 276 BP.
XX
XX AAL15054;
XX
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 7511.
XX
```

```
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
XX Homo sapiens.
OS
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US000798.
XX
XX 14-JAN-2000; 2000US-0176077P.
XX 14-MAR-2000; 2000US-0189167P.
XX 24-MAR-2000; 2000US-0192099P.
XX 29-MAR-2000; 2000US-0193480P.
XX 15-MAY-2000; 2000US-0205230P.
XX 09-JUN-2000; 2000US-0211315P.
XX 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX Claim 1; Page 1351; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
XX Sequence 276 BP; 15 A; 121 C; 30 G; 96 T; 0 U; 14 Other;
SQ
Query Match 80.0%; Score 14.4; DB 4; Length 276;
Best Local Similarity 88.2%; Pred. No. 3.2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCCCCTCTTGTGATAC 18
Db 38 CCCCCTCTCTGTTWC 54
RESULT 65
AAL191862
ID AAL191862 standard; cDNA; 307 BP.
XX
XX AAL191862;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 11922.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
```

PR 18-MAY-2000; 2000US-00577409.  
XX  
XX (HYSB-) HYSBQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-514838/56.  
XX  
XX P-PSDB; AAO119311.  
XX  
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
XX PT and treating e.g. leukemia, inflammation and immune disorders.  
XX  
XX  
PS Claim 1; SEQ ID NO 11922; 1399pp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AA179941-AA193841) and  
XX CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to  
XX CC cytokine, cell proliferation or cell differentiation or which may induce  
XX CC production of other cytokines in other cell populations. The  
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX CC peptide therapy. The polypeptides have various cytokine-like activities,  
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating  
XX CC activity, tissue growth factor activity, immunomodulatory activity and  
XX CC activin/inhibin activity and may be useful in the diagnosis and/or  
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX CC inflammation. Note: The sequence data for this patent did not form part  
XX CC of the printed specification, but was obtained in electronic format  
XX CC directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences  
SQ Sequence 307 BP; 48 A; 101 C; 69 G; 89 T; 0 U; 0 Other;

Query Match 80.0%; Score 14.4; DB 4; Length 307;  
Best Local Similarity 93.8%; Pred. No. 3.2e+03;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CCCCCCTCTTCGAT 16  
DB 101 CCCCCCTCTTCGAT 116

RESULT 66  
ABQ56427  
ID ABQ56427 standard; cDNA; 370 BP.  
XX  
XX ABQ56427;  
XX AC  
XX DT 02-AUG-2002 (first entry)  
XX  
XX Human colon cancer related nucleotide sequence SEQ ID NO:122.  
XX DE  
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
XX KW genetic analysis; diagnostic; antisense therapy; gene; ss.  
XX  
XX Homo sapiens.  
XX OS  
XX MO200229086-A2.  
XX PN  
XX 11-APR-2002.  
XX PD  
XX 02-OCT-2001; 2001WO-US030732.  
XX PF  
XX 02-OCT-2000; 2000US-0237271P.  
XX PR  
XX (PARB ) BAYER CORP.  
XX PA  
XX Burgess C, Asle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;  
XX PI Thaglingam A, Lewis ME;  
XX  
XX WPI; 2002-426115/45.  
XX  
XX New isolated nucleic acid that is differentially expressed in cancer  
XX PT tissues useful for determining the presence of colon cancer in a cell or  
XX PT tissue type, and in antisense therapy.

PS Claim 1; Fig 1; 796pp; English.  
XX  
XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
XX CC expressed in cancer tissues, ABB78993 to ABB79004 represent proteins  
XX CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
XX CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
XX CC encoded by (I) is useful for detecting cancer in a patient sample, and  
XX CC for detecting the presence or absence of a polynucleotide encoded by a  
XX CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived  
XX CC from (I) can be used for determining the presence of a nucleic acid which  
XX CC hybridizes to (I), and for determining the phenotype of cells in a sample  
XX CC of cells from a patient. (I) is useful for determining the presence of  
XX CC colon cancer in a cell or tissue type, for determining the presence or  
XX CC state of other type of cancer, in antisense therapy, to generate  
XX CC macroarrays on a solid surface, to identify a chromosome on which the  
XX CC corresponding gene resides, and in tissue profiling, forensics, genetic  
XX CC analysis, mapping and diagnostic applications. (I) can be used to raise  
XX CC antibodies, and to screen for peptide analogues and antagonists  
SQ Sequence 370 BP; 112 A; 77 C; 82 G; 91 T; 0 U; 8 Other;

Query Match 80.0%; Score 14.4; DB 6; Length 370;  
Best Local Similarity 93.8%; Pred. No. 3.3e+03;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CCCCCCTCTTCGAT 16  
DB 15 CCCCCCTCTTCGAT 30

RESULT 67  
ABN16125  
ID ABN16125 standard; cDNA; 378 BP.  
XX  
XX ABN16125;  
XX AC  
XX DT 24-JUN-2002 (first entry)  
XX  
XX Human ORFX polynucleotide sequence SEQ ID NO:727.  
XX DE  
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
XX KW hyperproliferative disorder; psoriasis; benign tumor; hemorrhage;  
XX KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
XX KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
XX KW hypertension; hypothyroidism; cholesterol ester storage disease;  
XX KW immune deficiency; immune disorder; infectious disease;  
XX KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
XX KW myasthenia gravis; gene; ss.  
XX  
XX Homo sapiens.  
XX OS  
XX MO200192523-A2.  
XX PN  
XX 06-DEC-2001.  
XX PD  
XX 29-MAY-2001; 2001WO-US010836.  
XX PF  
XX 30-MAY-2000; 2000US-0206132P.  
XX PR  
XX 29-AUG-2000; 2000US-0228716P.  
XX PA  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shinkets RA, Leach MD;  
XX PI  
XX WPI; 2002-106308/14.  
XX DR  
XX P-PSDB; ABP00373.  
XX  
XX Novel human polypeptides and polynucleotides useful for diagnosing,  
XX PT preventing and treating cardiovascular disease, neurodegenerative,  
XX PT hyperproliferative disorders and autoimmune disorders.  
XX  
XX Disclosure, SEQ ID NO 727; 1037pp; English.

CC The present invention describes substantially purified human proteins  
CC (referred to as open reading frame, ORFX, where X is 1-1191 (see Table 1  
CC in the specification). ABR15762 to ABR27252 encode the human ORFX  
CC proteins given in ABR00010 to ABR11500. ORFX proteins are useful for  
CC treating or preventing a pathology associated with an ORFX-associated  
CC disorder in humans, and in the manufacture of a medicament for treating a  
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
CC sequences can be used in gene therapy. ORFX sequences can be used in the  
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage diseases, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage. N.B. The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at [ftp.wipo.int/pub/published\\_poc\\_sequences](http://wipo.int/pub/published_poc_sequences)  
CC  
XX  
SQ Sequence 378 BP; 53 A; 125 C; 110 G; 90 T; 0 U; 0 Other;  
Query Match 80.0%; Score 14.4; DB 6; Length 378;  
Best Local Similarity 93.8%; Pred. No. 3.3e+03;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
DY 2 CCCCCTCTCTGATA 17  
45 CCCCCTCTCTGATA 60  
DB  
RESULT 68  
ABX64829  
ID ABX64829 standard; cDNA; 393 BP.  
XX  
AC ABX64829;  
XX  
DT 03-MAR-2003 (first entry)  
XX  
DE Human gene trapped sequence (GTS) cDNA SEQ ID NO 436.  
XX  
KW Human; gene trapped sequence; GTS; gene discovery; forensic; mapping;  
KW gene therapy; antisense therapy; prenatal analysis; autoimmune disease;  
KW developmental disorder; aging; cancer; Crohn's disease; schizophrenia;  
KW multiple sclerosis; skin disorder; rheumatoid arthritis; skeletal system;  
KW atherosclerosis; cardiovascular disease; degenerative disease; neural;  
KW Alzheimer's disease; osteoporosis; asthma; infection; ss.  
XX  
XX Homo sapiens.  
XX  
PN US2002110809-A1.  
XX  
PD 15-AUG-2002.  
XX  
PF 28-APR-2000; 2000US-00560863.  
XX  
PR 30-APR-1999; 99US-0132408P.  
XX  
PA (NEHL/) NEHLS M. C.  
PA (ZAMB/) ZAMEROWICZ B.  
PA (SAND/) SANDS A. T.  
XX  
PI Nehls MC, Zambrowicz B, Sands AT;  
XX  
DR WPI; 2003-090170/08.  
XX  
PT Novel human polynucleotides that correspond to human gene trapped  
PT sequences useful for gene discovery, gene therapy, as markers for gene

PT expression analysis, forensic analysis and determining genetic basis of  
PT diseases.  
XX  
PS Claim 1; SEQ ID NO 436; 36pp; English.  
XX  
XX This invention describes isolated or purified polynucleotides  
CC corresponding to human gene trapped sequence (GTS) comprising a sense or  
CC antisense sequence chosen from 1000 sequences or its fragment of 8  
CC contiguous nucleotides. GTSs are useful for gene discovery and as markers  
CC for gene expression analysis, identifying and mapping the coding regions  
CC of the mammalian and particularly human, genome, forensic analysis, for  
CC cross species hybridisation analysis, genetic manipulation, antisense  
CC inhibition, gene targeting, identification or generation of full-length  
CC cDNA, mapping the human genome, gene or antisense therapy, gene delivery  
CC and determining the genetic basis of human disease. Portion of the GTS  
CC sequences are useful as a hybridisation probe or for chromosome mapping,  
CC and can be incorporated into phase display system that can be used to  
CC screen for proteins, or other ligands, that are capable of binding an  
CC amino acid sequence encoded by the GTS sequences. The GTS sequences are  
CC also useful to regulate gene expression, as a part of ribozyme and/or  
CC triple helix sequences that can be used to regulate gene expression, as  
CC components of diagnostic methods, for analysing single nucleotide  
CC polymorphisms, and also as genetic markers for prenatal analysis of  
CC congenital traits or defects. The polynucleotides of the invention are  
CC useful for diagnosis, prognosis of disorders involving developmental and  
CC differentiation processes and for the identification of subjects having a  
CC predisposition to such disorders. Diseases or natural processes that can  
CC be correlated with the expression of mutant or normal, variants of GTSs  
CC include, aging, cancer, autoimmune disease, Crohn's disease, multiple  
CC sclerosis, immune disorders, cardiovascular disease, degenerative diseases  
CC arthritis, atherosclerosis, cardiovascular disease, osteoporosis,  
CC of the neural or skeletal systems, Alzheimer's disease, osteoporosis,  
CC asthma, and infections. GTSs can also be used to identify the specific  
CC locations of exon splice junctions, which are important in the study of  
CC disease and cancer. Modulating the level of expression of one or more  
CC genes and/or regulating activity of one or more peptides or proteins is  
CC useful for modifying development and cell differentiation disorders. ABX4402-ABX65401  
CC development and cell differentiation disorders. ABX4402-ABX65401  
CC represent the human gene trapped cDNA sequences described in the  
CC disclosure of the invention. NOTE: The sequence data for this patent did  
CC not form part of the printed specification but was obtained in electronic  
CC format directly from USPTO at  
CC [seqdata.uspto.gov/sequence.html?docid=20020110809](http://seqdata.uspto.gov/sequence.html?docid=20020110809)  
XX  
SQ Sequence 393 BP; 103 A; 119 C; 91 G; 80 T; 0 U; 0 Other;  
Query Match 80.0%; Score 14.4; DB 8; Length 393;  
Best Local Similarity 93.8%; Pred. No. 3.3e+03;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
DY 2 CCCCCTCTCTGATA 17  
152 CCCCCTCTCTGACA 167  
DB  
RESULT 69  
ADF59862/c  
ID ADF59862 standard; cDNA; 406 BP.  
XX  
AC ADF59862;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Human contig polynucleotide sequence SEQ ID NO:2229.  
XX  
KW biological activity; genetic engineering; hybridisation probe; oligomer;  
KW primer; chromosome mapping; gene mapping; recombinant protein production;  
KW human; gene; ss.  
XX  
XX Homo sapiens.  
XX  
OS  
PN WO2003080795-A2.  
XX

PD 02-OCT-2003.  
 XX 09-AUG-2002; 2002WO-US025485.  
 PF 09-AUG-2001; 2001US-0311261P.  
 PR  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YT, Yang Y, Wang Z, Weng G, Ma Y;  
 PI WPI; 2003-876918/81.  
 DR P-PSDB; ADF60314.  
 XX  
 XX New polynucleotides, useful as hybridization probes, oligomers or  
 PT primers, for chromosome or gene mapping, for the recombinant production  
 of proteins, and for generating antisense DNA or RNA.  
 PS Example 2; SEQ ID NO 2229; 571pp; English.  
 CC The present invention describes isolated polynucleotide sequences (I),  
 CC which encode polypeptides (II) with biological activity. Also described:  
 CC (1) a vector comprising (I); (2) an expression vector comprising (I); (3)  
 CC a host cell genetically engineered to comprise (I) which is operatively  
 CC associated with a regulatory sequence that modulates expression of (I) in  
 CC the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition  
 CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed  
 CC against the polypeptide of (4); (7) detecting (I) or the polypeptide of  
 CC (4) in a sample; (8) identifying a compound that binds to the polypeptide  
 CC of (4); (9) producing the polypeptide of (4); and (10) a collection of  
 CC polynucleotides comprising at least one of the polynucleotide sequences  
 CC (I). The polynucleotides (I) can be used as hybridization probes,  
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant  
 CC production of proteins, and for generating antisense DNA or RNA. The  
 CC present sequence represents a human contig polynucleotide sequence, which  
 CC is used in an example from the present invention.  
 XX  
 SQ Sequence 406 BP; 83 A; 125 C; 141 G; 57 T; 0 U; 0 Other;  
 Query Match 80.0%; Score 14.4; DB 10; Length 406;  
 Best Local Similarity 93.8%; Pred. No. 3.3e+03;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CCCCCCTCTTGAT 16  
 DB 234 CCCCCCTCTTGAT 219  
 RESULT 70  
 ABL80291  
 ID ABL80291 standard; cDNA; 419 BP.  
 AC ABL80291;  
 XX  
 XX 17-MAY-2002 (first entry)  
 DT  
 XX  
 DE Human ovarian cancer related cDNA clone SEQ ID NO:3269.  
 XX  
 KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200192581-A2.  
 PD 06-DEC-2001.  
 XX  
 PF 29-MAY-2001; 2001WO-US017756.  
 XX  
 PR 26-MAY-2000; 2000US-0207484P.  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Algate PA, Harlocker SL, Jones R;  
 XX

DR WPI; 2002-122075/16.  
 XX  
 XX Composition for therapy and diagnosis of ovarian cancer comprising  
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
 PT polypeptide, antibody specific to polypeptide or T cell expressing  
 PT polypeptide.  
 PS Claim 1; SEQ ID NO 3269; 489pp; English.  
 XX  
 XX The present invention describes a composition (I) comprising: carriers  
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour  
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)  
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,  
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),  
 CC or antigen presenting cells that express (II). (I) has cytostatic  
 CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for  
 CC detecting ovarian cancer in a patient's biological sample preferably  
 CC serum or ovarian tissue. The method comprises contacting a biological  
 CC sample from a patient with (IV), detecting the amount of polynucleotide  
 CC hybridising to (IV) and comparing the amount to a predetermined cutoff  
 CC value and thereby detecting ovarian cancer in the patient, where the  
 CC amount of polynucleotide hybridising to (IV) is detected preferably by  
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is  
 CC useful for stimulating and/or expanding T cells specific for an ovarian  
 CC tumour protein comprising contacting T cells with (III) or (II). (III) is  
 CC useful in design and preparation of ribozyme molecules for inhibiting  
 CC expression of the tumour polypeptides and proteins in tumour cells; and  
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA  
 CC library using well known techniques  
 XX  
 SQ Sequence 419 BP; 76 A; 114 C; 97 G; 132 T; 0 U; 0 Other;  
 Query Match 80.0%; Score 14.4; DB 6; Length 419;  
 Best Local Similarity 93.8%; Pred. No. 3.3e+03;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 CCCCCCTCTTGATAC 18  
 DB 107 CCCCCCTCTTGATAC 122  
 Search completed: May 6, 2006, 19:49:43  
 Job time : 301 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 5, 2006, 14:36:34 ; Search time 2331 Seconds  
(without alignments)  
361.290 Million cell updates/sec

Title: US-10-659-980A-5

Perfect score: 18

Sequence: 1 cccccctctctgcatac 18

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 70 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_hic: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_est7: \*  
9: gb\_g881: \*  
10: gb\_g882: \*  
11: gb\_g883: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.4	91.1	251	9	B01685
2	16.4	91.1	318	7	CK764583
3	16.4	91.1	361	2	BE611052
4	16.4	91.1	528	9	BH871872
5	16.4	91.1	544	3	BJ092766
6	16.4	91.1	561	9	CC763441
7	16.4	91.1	565	3	B1839321
8	16.4	91.1	566	5	B0871803
9	16.4	91.1	584	9	A2711917
10	16.4	91.1	615	8	CE036382
11	16.4	91.1	626	8	CK029731
12	16.4	91.1	631	10	CE354701
13	16.4	91.1	638	8	DN267386
14	16.4	91.1	649	5	BY755328
15	16.4	91.1	708	3	B619128
16	16.4	91.1	714	9	CC633555
17	16.4	91.1	725	10	CG191479
18	16.4	91.1	736	11	CR104832
19	16.4	91.1	749	9	BZ980674
20	16.4	91.1	779	10	CG109375
21	16.4	91.1	782	10	AG421087
22	16.4	91.1	837	2	BE961386

#### ALIGNMENTS

23	16.4	91.1	857	8	DN789757
24	16.4	91.1	892	10	AG344768
25	16.4	91.1	1044	10	AG311991
26	16.4	91.1	1093	3	BM478544
27	16.4	91.1	1186	7	CK231738
28	16.4	91.1	1188	9	CC317382
29	16.4	91.1	1300	9	BM809604
30	16.4	91.1	1515	3	BM475412
31	16.4	88.9	480	10	CZ179143
32	16.4	88.9	536	10	CG585619
33	16.4	88.9	556	9	AQ551912
34	16.4	88.9	582	3	BP292366
35	16.4	88.9	1000	2	BG027192
36	16.4	88.9	1278	5	BU149766
37	15.4	85.6	178	8	DN104238
38	15.4	85.6	179	2	BE142033
39	15.4	85.6	192	9	AZ746799
40	15.4	85.6	201	1	BM160827
41	15.4	85.6	262	7	CK818409
42	15.4	85.6	271	9	CE076988
43	15.4	85.6	282	2	BB399270
44	15.4	85.6	304	2	BB396052
45	15.4	85.6	306	3	BI683470
46	15.4	85.6	313	1	AW883754
47	15.4	85.6	325	2	BB242660
48	15.4	85.6	328	2	BB316251
49	15.4	85.6	328	9	AZ722361
50	15.4	85.6	329	1	AM482062
51	15.4	85.6	332	3	BJ185054
52	15.4	85.6	356	7	CO878314
53	15.4	85.6	379	7	CN439166
54	15.4	85.6	381	6	CB423611
55	15.4	85.6	384	4	AF227818
56	15.4	85.6	386	2	BF544009
57	15.4	85.6	387	7	CK443748
58	15.4	85.6	387	7	CE777815
59	15.4	85.6	387	10	BY472844
60	15.4	85.6	388	5	BY702625
61	15.4	85.6	399	5	BM124901
62	15.4	85.6	400	7	CK443747
63	15.4	85.6	405	7	BJ197399
64	15.4	85.6	415	2	BF761320
65	15.4	85.6	415	2	BU201372
66	15.4	85.6	416	3	BU201372
67	15.4	85.6	419	5	BU917350
68	15.4	85.6	419	1	AV667682
69	15.4	85.6	432	9	AZ241055
70	15.4	85.6	432	9	AZ241055

RESULT 1  
B01685  
LOCUS  
DEFINITION  
csrcL-136b5-u csrcL flow sorted Chromosome 11 specific cosmid Homo  
sapiens genomic clone csrcL-136b5, genomic survey sequence.

ACCESSION  
B01685  
VERSION  
B01685.1 GI:1410963  
KEYWORDS  
GSS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE  
1 (bases 1 to 251)  
Evans,G.A., Burdick,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M.,  
Jones,D., Ward,T., Gilligan,B., Schlegelmann,J., Probst,S.,  
Hartley,J., DeFord,J., McFarland,J., Butzinski,K., Khan,M.,  
Kupfer,K. and Garner,H.R.  
Genomic Sequence Sampled Map of Chromosome 11  
Unpublished (1996)

## COMMENT

Contact: Evans GA, Shane Probst  
McDermott Center for Human Growth and Development  
University of Texas Southwestern Medical Center At Dallas  
5323 Harry Hines Blvd, Dallas TX 75235-8591  
Tel: 214-648-1600  
Fax: 214-648-1666  
Email: g.evans@utsw.swmed.edu, shane@mcdermott.swmed.edu

## PCR PRIMERS

FORWARD: TGTGATCTATCCAGAGCC

BACKWARD: ACACGATTGAAGTCACAGTC

Seq primer: T7

Class: cosmid ends

High quality sequence stop: 251.

## FEATURES

## source

Location/Qualifiers

1..251

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="CSRL-136b5"

/sex="female"

/cell\_type="chimeric hamster somatic cell hybrid"

/clone\_1ib="CSRL flow sorted Chromosome 11 specific cosmid"

/note="Vector: sCos-1; Human Chromosome 11 specific cosmid library prepared from flow sorted human Chromosome 11 derived from Chinese Hamster Ovary (CHO) monochromosomal somatic cell hybrid, J1"

## ORIGIN

## Query Match

Best Local Similarity 91.1%; Score 16.4; DB 9; Length 251;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTCTTGATAC 18

Db 170 CCCCCCTCTTGATAC 187

RESULT 2 CK764583 318 bp mRNA linear EST 09-JUN-2005

LOCUS 1tu01-13m93-g03 318 bp mRNA linear EST 09-JUN-2005

DEFINITION 1tu01-13m93-g03 5', mRNA sequence.

ACCESSION CK764583

VERSION CK764583.1 GI:42655034

KEYWORDS EST.

SOURCE Liriodendron tulipifera

ORGANISM Liriodendron tulipifera

Bukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Magnoliales;

Magnoliales; Liriodendron.

1 (bases 1 to 318)

dePamphilis,C., Carlson,J., Ma,H., Soltis,D., Soltis,P.,

Oppenheimer,D., Frolich,M., Doyle,J., Tanksey,S., Webb,M.,

Leebens-Mack,J., Landherr,L., Schlarbaum,S., Ilut,D. and Wall,K.

Generation of ESTs from early flower buds of Liriodendron

tulipifera

Unpublished (2003)

Contact: Claude dePamphilis or James Leebens-Mack

Penn State University

208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn

State University, University Park, PA 16802, USA

Tel: 814 863 6413

Fax: 814 865 9131

Email: cwid@psu.edu or jh110@psu.edu

The sequence provided is trimmed of vector and low quality regions.

Full sequence and original trace file are available from the Plant

Genome Network website (http://pgn.cornell.edu)

Plate: 1tu01-13m93 row: 9 column: 03

Seq primer: M13F.

Location/Qualifiers

1..318

FEATURES

source

## ORIGIN

## Query Match

Best Local Similarity 91.1%; Score 16.4; DB 7; Length 318;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTCTTGATAC 18

Db 221 CCCCCCTCTTGATAC 204

RESULT 3 BE611052 361 bp mRNA linear EST 13-JUL-2004

LOCUS sg74c02.y1 Gm-cl048 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

Gm-cl048-1179.5, similar to TR:040512 Q40512 PHOTOSYSTEM I

LIGHT-HARVESTING CHLOROPHYLL A/B-BINDING PROTEIN. ;, mRNA sequence.

BE611052

VERSION BE611052.1 GI:9902084

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 361)

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corryell,V.,

Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,

Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,

Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,

Schurk,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,

McCaun,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: esewatson.wustl.edu

When it has been determined, an EST from the other end of this

clone is listed in the 'Other ESTs on clone' field. Trace

considered overall poor quality This clone is available through:

Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone:

800 423 4163; email: info@biogeneticservices.com)

Insert Length: 940 Std Error: 0.00

High quality sequence stop: 1.

Location/Qualifiers

1..361

/organism="Glycine max"

/mol\_type="mRNA"

/cultivar="Clark"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl048-1179"

/tissue\_type="whole seedlings of greenhouse grown plants"

/dev\_stage="1 week old"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The Clark NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shomaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from whole seedlings of 1 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). The library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 361;  
 Best Local Similarity 94.4%; Pred. No. 2.7e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18  
 |||||  
 288 CCCCCCTCTTCGTGATAC 271

RESULT 4  
 BH871872 528 bp DNA linear GSS 05-AUG-2002  
 LOCUS hm37c05\_g1 WGS-Zmayaf (JMI07 adapted methyl filtered) Zea mays  
 DEFINITION genomic clone hm37c05 5', genomic survey sequence.  
 ACCESSION BH871872  
 VERSION BH871872.1 GI:22107769  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 528)  
 Rabinowicz,P.D., O'Shaughnessy,A.L., Ballig,V., Dedhia,N., Katzenburger,F., King,L., Miller,B., Miller,S., Nascimeto,L., Zlatavern,T., McCombie,W.R. and Martensen,R.A.  
 Genomic shotgun sequences from Zea mays (methyl-filtered) unpublished (2002)  
 Contact: W. Richard McCombie  
 Lita Annenberg Hazen Genome Sequencing Center  
 Cold Spring Harbor Laboratory  
 PO Box 100, Cold Spring Harbor, NY 11724, USA  
 Tel: 516 367 8884  
 Fax: 516 367 8874  
 Email: mcombie@cshl.org  
 Plate: hm37 row: C Column: 05  
 Seq primer: -21M3UnivRev  
 Class: shotgun  
 High quality sequence stop: 528.  
 Location/Qualifiers  
 1..528

FEATURES  
 source

/organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /lab\_host="JMI07 or DH5a"  
 /clone="hm37c05"  
 /note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were

between 0.8 and 3 kb and were cloned into the vector (x/y reads in M13mp19, b/g reads in pUC19). The same ligation was transformed in either JMI07 or DH5a. "

## ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 528;  
 Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18  
 |||||  
 Db 107 CCCCCCTCTTCGTGATTC 90

RESULT 5  
 BU092766 544 bp mRNA linear EST 01-OCT-2003  
 LOCUS BU092766 NIBB Mochii normalized Xenopus tailbud library Xenopus  
 DEFINITION laevis cDNA clone XL057f10 3', mRNA sequence.

ACCESSION BU092766  
 VERSION BU092766.1 GI:17592031  
 KEYWORDS EST.  
 SOURCE Xenopus laevis (African clawed frog)  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Xenopus.  
 1 (bases 1 to 544)  
 Kitayama,A., Teraoka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara,Y.  
 Expressed genes in X. laevis embryo  
 Unpublished (2001)  
 Contact: Tadashi Shin-i  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp  
 The information of this clone is available through the following URL.  
 http://xenopus.nibb.ac.jp.  
 Location/Qualifiers  
 1..544

FEATURES  
 source

/organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="XL057f10"  
 /tissue\_type="whole embryo"  
 /dev\_stage="stage 25"  
 /clone\_lib="NIBB Mochii normalized Xenopus tailbud library"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 3; Length 544;  
 Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGTGATAC 18  
 |||||  
 Db 265 CCCCCCTCTTCGTGACAC 282

## RESULT 6

LOCUS CC763441 561 bp DNA linear GSS 27-JUN-2003  
 DEFINITION CH240\_4C2.TJ CHORI-240 Bos taurus genomic clone CH240\_4C2, genomic survey sequence.

ACCESSION CC763441  
 VERSION CC763441.1 GI:32309939  
 KEYWORDS GSS.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus





TITLE  
Lettuce and Sunflower ESTs from the Compositae Genome Project  
JOURNAL  
http://comgenomics.ucdavis.edu/  
COMMENT  
Unpublished (2002)  
Contact: Alexander Kozik [R.W.Michelmore]  
Department of Vegetable Crops, R.W.Michelmore Lab  
University of California at Davis (UCD)  
Assumundson Hall, UCD, Davis, CA 95616, USA  
Tel: 1-(530)-742-1742  
Fax: 1-(530)-752-9659  
Email: akozik@ucdavis.edu [michelmoreveg@mail.ucdavis.edu]  
belongs to contig QG\_CA\_Contig4405, see http://cgdb.ucdavis.edu/  
for details.  
Plate: QG12 row: N column: 19.  
Location/Qualifiers

FEATURES  
source  
1..566  
/organism="Lactuca sativa"  
/mol\_type="mRNA"  
/cultivar="Salinas"  
/db\_xref="taxon:4236"  
/clone="QG12N19"  
/lab\_host="E.coli"  
/note="Vector: pBRCDNA5f1AB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/  
TAG LIB=QG ABCDI lettuce salinas  
TAG\_SEQ=CGTTGACGGG"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 5; Length 566;  
Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 CCCCCCTCTTGTGATAC 18  
|||||  
543 CCCCCCTCTTGTGATAC 526

## ORIGIN

Query Match 91.1%; Score 16.4; DB 5; Length 566;  
Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGTGATAC 18  
|||||  
Db 543 CCCCCCTCTTGTGATAC 526

## RESULT 9

AZ711917/c 584 bp DNA linear GSS 24-JAN-2001  
LOCUS  
DEFINITION  
RPCI-24-129013.TV RPCI-24 Mus musculus genomic clone  
RPCI-24-129013, genomic survey sequence.  
ACCESSION  
VERSION  
A2711917.1 GI:12445119  
KEYWORDS  
SOURCE  
ORGANISM  
GSS.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Mus.  
1 (bases 1 to 584)  
Zhao, S., Nierman, W., Malek, J., Shatman, S., Akinc, B., Levine, M.,  
Tegay, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregiorgis, E.,  
Russell, D., de Jong, P. and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-24  
Unpublished (1999)  
Other\_GSSs: RPCI-24-129013.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: schao@tigr.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

library availability, please contact Pieter de Jong  
(pdjong@mail.cho.org). Clones may be purchased from BACPAC  
Resources (http://www.choi.org/bacpac/orderingframe.html). BAC end  
page: http://www.tigr.org/cdb/bac\_ends/mouse/bac\_end\_intro.html  
Plate: 129 row: O column: 13  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers

FEATURES  
source  
1..584  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-129013"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/clone\_lib="RPCI-24"  
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI;  
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The  
library was cloned in the pTARBAC1 cloning vector at the  
BamHI sites using MboI partially digested male C57BL/6J  
DNA."

ORIGIN  
Query Match 91.1%; Score 16.4; DB 9; Length 584;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 CCCCCCTCTTGTGATAC 18  
|||||  
Db 350 CCCCCCTCTTGTGATAC 333

## ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 584;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGTGATAC 18  
|||||  
Db 350 CCCCCCTCTTGTGATAC 333

## RESULT 10

CE036382 616 bp DNA linear GSS 24-SEP-2003  
LOCUS  
DEFINITION  
tigr-gss-dog-17000348575092 Dog library Canis familiaris genomic,  
genomic survey sequence.  
ACCESSION  
VERSION  
CE036382.1 GI:35064065  
KEYWORDS  
SOURCE  
ORGANISM  
GSS.  
Canis familiaris (dog)  
Canis familiaris  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;  
Canis.  
1 (bases 1 to 616)  
Kirkness, E.F., Batra, V., Halpern, A.L., Levy, S., Remington, K.,  
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
Venter, J.C.  
The dog genome: survey sequencing and comparative analysis  
Science 301 (5641), 1898-1903 (2003)  
1451627  
Contact: Kirkness EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirknes@tigr.org  
Class: shotgun.  
Location/Qualifiers

FEATURES  
source  
1..616  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/note="Site 1: BstXI; Libraries were prepared from  
peripheral blood"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 616;

Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18  
|||||  
Db 261 CCCCCCTTCTGGATCC 278

## RESULT 11

LOCUS CX029731 625 bp mRNA linear EST 03-JAN-2005  
DEFINITION 1341670 NCCCW 10RT#3 Oncorhynchus mykiss cDNA 5', mRNA sequence.

ACCESSION CX029731  
VERSION CX029731.1 GI:56974085

KEYWORDS EST.  
SOURCE Oncorhynchus mykiss (rainbow trout)  
ORGANISM Oncorhynchus mykiss

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

REFERENCE 1 (bases 1 to 625)  
Yao, J., Gahr, S., and Rexroad, C.E.  
10RT#3 egg, NCCCW/WVU EST Project, Phase II

JOURNAL Unpublished (2004)  
COMMENT Contact: Rexroad CE  
USDA, ARS, National Center for Cool and Cold Water Aquaculture  
11876 Leetown Road, Kearneysville, WV 25430, USA  
Tel: 304 724 8340 x2129  
Fax: 304 725 0351  
Email: crexroad@ncccw.ars.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and  
trimmed with the aid of the trim\_alt option. Vector identified with  
cross\_match v0.990329.  
Plate: 115 row: L column: 4  
Seq primer: GYATACGACTCCTACTATAGGG.

FEATURES  
source Location/Qualifiers

1..625  
/organism="Oncorhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/lab\_host="DH10B"  
/clone\_lib="NCCCW 10RT#3"  
/note="Vector: pCMV Sport6.0; WVU oocyte library RT-Egg."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 8; Length 625;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18  
|||||  
Db 563 CCCCCCTTATGATAC 580

## RESULT 12

LOCUS CE354701 631 bp DNA linear GSS 26-SEP-2003  
DEFINITION t1gr-gss-dog-17000361484864 Dog Library Canis familiaris genomic,  
genomic survey sequence.

ACCESSION CE354701  
VERSION CE354701.1 GI:36192654

KEYWORDS GSS.  
SOURCE Canis familiaris (dog)  
ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;  
Canis.

REFERENCE 1 (bases 1 to 631)  
Kirkness, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K.,  
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Frazer, C.M. and  
Venter, J.C.

TITLE The dog genome: survey sequencing and comparative analysis  
JOURNAL Science 301 (5641), 1898-1903 (2003)

PUBMED 14512627  
Contact: Kirkness EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirkness@tigr.org

Class: shotgun.  
Location/Qualifiers

## FEATURES

source 1..631  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/note="Site 1: BactX; Libraries were prepared from  
peripheral blood"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 631;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18  
|||||  
Db 188 CCCCCCTTCTGGATCC 205

## RESULT 13

LOCUS DN267386 638 bp mRNA linear EST 02-MAR-2005  
DEFINITION L1B30320\_018\_G05\_Sp6\_1 L1B30320 Canis familiaris cDNA clone  
L1B30320\_018\_G05, mRNA sequence.

ACCESSION DN267386  
VERSION DN267386.1 GI:60435996

KEYWORDS EST.  
SOURCE Canis familiaris (dog)  
ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;  
Canis.

REFERENCE 1 (bases 1 to 638)  
Staten, N.R.

AUTHORS Direct Submission (Staten, N.R.)  
TITLE Unpublished (2005)  
JOURNAL Contact: Nick Staten  
COMMENT Tel: 636 247 6855  
Email: nicholas.r.staten@pfizer.com.

FEATURES  
source Location/Qualifiers

1..638  
/organism="Canis familiaris"  
/mol\_type="mRNA"  
/db\_xref="taxon:9615"  
/clone="L1B30320\_018\_G05"  
/issue\_type="lymph node"  
/lab\_host="DH10B"  
/clone\_lib="L1B30320"  
/note="Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: SalI;  
barcode 459"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 8; Length 638;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18  
|||||  
Db 238 CCCCCCTTCTGGATCC 255

RESULT 14  
BY755362/c

LOCUS BY755362 649 bp mRNA linear EST 09-JUL-2003  
DEFINITION BY755362 RIKEN full-length enriched, embryo RCB-0549 C1e-H3 CDNA  
Mus musculus cDNA clone G431003018 3', mRNA sequence.  
ACCESSION BY755362  
VERSION BY755362.1 GI:27187758  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Murinae; Mus.  
1 (bases 1 to 649)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schombach, C., Cojocari, T., Baldarelli, R., Hill, D. P., Bull, C.,  
Hume, D. A., Quackenbush, J., Schriml, L. M., Kanpin, A., Matsuda, H.,  
Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Bruscia, V.,  
Chottha, C., Corbani, L. E., Cousins, S., Dalla, E., Dragan, T. A.,  
Fletcher, C. P., Forrest, A., Frazer, K. S., Gaasterland, T.,  
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R. M., Kling, B. L., Konagaya, A.,  
Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglocz, D. R.,  
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W. J., Petosa, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramchandran, S.,  
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carrincci, P.,  
Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,  
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
Aizawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,  
Rogers, J., Birney, E. and Hayashizaki, Y.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length CDNA  
Nature 420, 563-573 (2002)  
12466851  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carrincci, P.,  
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,  
Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,  
Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, N., Nakamura, M.,  
Nomura, K., Numazaki, R., Ohno, M., Osato, N., Saito, R., Sakazume, N.,  
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,  
Takeda, Y., Waki, K., Watanaka, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submissions  
Computational Analysis of Full-length Mouse CDNA Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected CDNA to  
prepare full-length CDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multichannel sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length CDNA  
nonredundant CDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
FEATURES  
source Location/Qualifiers  
1..649  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="G431003018"  
/cell\_line="RCB-0549 C1e-H3"  
/dev\_stage="embryo"  
/clone\_id="RIKEN full-length enriched, embryo RCB-0549  
C1e-H3 CDNA"  
ORIGIN  
Query Match 91.1%; Score 16.4; DB 5; Length 649;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CCCCCCTCTTGATAC 18  
|||||  
Db 197 CCCCCCTCTTGAAAC 180  
RESULT 15  
Bj619128/c 708 bp mRNA linear EST 01-OCT-2003  
LOCUS Bj619128 NIBB Mochii normalized Xenopus early gastrula library  
DEFINITION Xenopus laevis cDNA clone X118612 5', mRNA sequence.  
ACCESSION Bj619128  
VERSION Bj619128.1 GI:37257590  
KEYWORDS EST.  
SOURCE Xenopus laevis (African clawed frog)  
ORGANISM Xenopus laevis  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
Xenopodinae; Xenopus; Xenopus.  
1 (bases 1 to 708)  
Kikayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and  
Kohara, Y.  
Expressed genes in X. laevis embryo  
Unpublished (2001)  
Contact: Tadashi Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@gene.nig.ac.jp  
The information of this clone is available through the following  
URL.  
http://xenopus.nibb.ac.jp.  
FEATURES  
source Location/Qualifiers  
1..708  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="X118612"  
/tissue\_type="whole embryo"  
/dev\_stage="stage 10.5"  
/clone\_id="NIBB Mochii normalized Xenopus early gastrula  
library"  
ORIGIN  
Query Match 91.1%; Score 16.4; DB 3; Length 708;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CCCCCCTCTTGATAC 18  
|||||  
Db 523 CCCCCCTCTTGACAC 506

RESULT 16  
CC633555/c  
LOCUS  
DEFINITION CC633555 714 bp DNA linear GSS 19-JUN-2003  
OGUJ0577H ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0456117,  
genomic survey sequence.  
ACCESSION CC633555  
VERSION CC633555  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
AUTHORS Whiteblaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,  
Resnick,A., Fraser,C.M., Robidin,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)  
COMMENT  
Other\_GSSs: OGUJ0577V  
Contact: Cathy Whiteblaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteblaw@tigr.org  
Seq primer: TR  
Classes: methylation filtered.  
Location/Qualifiers  
1..714  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0456117"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="Vector: pBCSk-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 9; Length 714;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGATAC 18  
|||||  
69 CCCCCCTCTTCGATTC 52

RESULT 17  
CG191479  
LOCUS  
DEFINITION CG191479 725 bp DNA linear GSS 21-AUG-2003  
PUIF152TB ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTA0571108,  
genomic survey sequence.  
ACCESSION CG191479  
VERSION CG191479  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
AUTHORS Whiteblaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennetzen,J.  
Maize Genomics Consortium  
Unpublished (2003)  
Other\_GSSs: PUIF152TD  
Contact: Cathy Whiteblaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteblaw@tigr.org  
Seq primer: TR  
Classes: sheared ends.  
Location/Qualifiers  
1..725  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBTA0571108"  
/clone\_lib="ZM 0.6 1.0 KB"  
/note="Vector: pCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
cor selected genomic DNA library"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 10; Length 725;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGATAC 18  
|||||  
Db 193 CCCCCCTCTTCGATTC 210

RESULT 18  
CR104832  
LOCUS  
DEFINITION CR104832 736 bp DNA linear GSS 05-JUL-2004  
Forward strand read from insert in 3'HPT insertion targeting and  
chromosome engineering clone MHPPI51K02, genomic survey sequence.  
ACCESSION CR104832.1 GI:49852232  
VERSION  
KEYWORDS GSS; genome survey sequence; MICER.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridea; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
Rogers,J. and Bradley,A.  
Direct Submision  
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgehire,  
CB10 1SA, UK. http://www.sanger.ac.uk/MICER  
Location/Qualifiers  
1..736  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPPI51K02"  
/clone\_lib="MHP"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 11; Length 736;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTCGATAC 18  
|||||  
Db 719 CCTCCCTCTTCGATAC 736

RESULT 19  
BZ980674  
LOCUS  
DEFINITION BZ980674 749 bp DNA linear GSS 25-MAR-2003  
PUGIG30TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTa390F11,  
genomic survey sequence.  
ACCESSION BZ980674  
VERSION BZ980674.1 GI:29211710  
KEYWORDS GSS.  
SOURCE Zea mays

ORGANISM  
Zea mays  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 749)  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Benneetzen,J.  
TITLE  
Maize Genomics Consortium  
JOURNAL  
Unpublished (2003)  
COMMENT  
Other\_GSSs: PUTG024TB  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@cigr.org  
Seq primer: TF  
Class: sheared ends.  
Location/Qualifiers  
1..749  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZM0.6.1.0 KB"  
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
COT selected genomic DNA library"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 9; Length 749;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATAC 18  
|||||||  
Db 591 CCCCCCTCTTGATTC 608

RESULT 20  
CG109375 779 bp DNA linear GSS 20-AUG-2003  
LOCUS  
DEFINITION  
PUG024TB ZM 0.6.1.0 KB Zea mays genomic clone ZMMBTA0578D23,  
genomic survey sequence.  
CG109375  
ACCESSION  
CG109375.1 GI:33992812  
VERSION  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 779)  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Benneetzen,J.  
TITLE  
Maize Genomics Consortium  
JOURNAL  
Unpublished (2003)  
COMMENT  
Other\_GSSs: PUG024TD  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@cigr.org  
Seq primer: TR  
Class: sheared ends.  
Location/Qualifiers  
1..779  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"

FEATURES  
source

ORIGIN  
Query Match 91.1%; Score 16.4; DB 10; Length 779;  
Best Local Similarity 94.4%; Pred. No. 3e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATAC 18  
|||||||  
Db 711 CCCCCCTCTTGATTC 728

RESULT 21  
AG421087 782 bp DNA linear GSS 21-DEC-2004  
LOCUS  
DEFINITION  
Mus musculus molossinus DNA, clone:MSMG01-292F10.TJ, genomic survey  
sequence.  
AG421087  
AG421087.1 GI:48064150  
GSS.  
Mus musculus molossinus (Japanese wild mouse)  
Mus musculus molossinus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
1  
Abe,K., Noguchi,H., Tagawa,K., Yuzuriba,M., Toyoda,A., Kojima,T.,  
Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriaki,K. and  
Shiroishi,T.  
Contribution of Asian mouse subspecies Mus musculus molossinus to  
genomic constitution of strain C57BL/6J, as defined by BAC-end  
sequence-SNP analysis  
Genome Res. 14 (12), 2439-2447 (2004)  
15574823  
2 (bases 1 to 782)  
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
TITLE  
Direct Submission  
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suenho-chou,Tsukumi-ku, Yokohama, Kanagawa, 230-0045, Japan  
(E-mail:hattori@gsc.riken.jp; URL:http://hnp.gsc.riken.go.jp/,  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the mouse BAC library MSMG01. For BAC  
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).  
Tsukuba Institute, Bio Resource Center.  
The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
Koyadai, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@rtc.riken.jp  
PRIMERS  
Sequencing : TJ  
LIBRARY : pBACe3.6  
Vector : EcoRI  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI  
Location/Qualifiers  
1..782  
/organism="Mus musculus molossinus"  
/mol\_type="genomic DNA"  
/sub\_species="molossinus"  
/db\_xref="taxon:57486"  
/clone\_lib="MSMG01-292F10.TJ"  
/sex="male"  
/tissue="mixture of kidney and spleen"  
/clone\_lib="MSMG01 Mouse Male BAC Library"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 10; Length 782;  
Best Local Similarity 94.4%; Pred. No. 3e+03;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTTCTGGATAC 18  
 |||||  
 Db 379 CCCCCCTTTCTGGATAC 396

RESULT 22  
 BE961386/c 837 bp mRNA linear EST 04-OCT-2000  
 LOCUS 601648508R2 NIH\_MGC\_62 Homo sapiens cDNA clone IMAGE:3932400 3',  
 DEFINITION mRNA sequence.  
 BE961386  
 BE961386.1 GI:10572091  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
 1 (bases 1 to 837)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-rc@mail.nih.gov  
 Tissue Procurement: ATCC/DCTD/DTF  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L1CM769 row: P column: 01.  
 Location/Qualifiers  
 1..837  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3932400"  
 /tissue\_type="melanotic melanoma, high MDR"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 62"  
 /note="Organ: skin; Vector: pDNR-LIB (Clontech); Site\_1:  
 Site (ggcgccttgagcc); Site\_2: Site (ggcctatggcc);  
 Double-stranded cDNA was prepared from cell line RNA. 5'  
 and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CACGCGCATTAAGCC-3' and 3' adaptor  
 sequence: 5'-ATTCTAGAGCGGAGCGCGCAGCATG-dT(30)BN-3'  
 (where B = A, C, or G and N = A, C, G, or T). Average  
 insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 837;  
 Best Local Similarity 94.4%; Pred. No. 3e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTTCTGGATAC 18  
 |||||  
 Db 394 CCCCCCTTTCTGGATAC 377

RESULT 23  
 DN789757 857 bp mRNA linear EST 07-APR-2005  
 LOCUS 94242891 Sea Urchin primary mesenchyme cell cDNA library  
 DEFINITION Strongylocentrotus purpuratus cDNA clone FMCSPR2-174C5 5', mRNA  
 sequence.  
 DN789757  
 DN789757.1 GI:62379824

ACCESSION  
 VERSION

KEYWORDS  
 SOURCE  
 ORGANISM  
 EST.  
 Strongylocentrotus purpuratus  
 Strongylocentrotus purpuratus  
 Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 Echinoidea; Euechinoidea; Echinacea; Echinoidea;  
 Strongylocentrotidae; Strongylocentrotus.  
 1 (bases 1 to 857)  
 Zhu X., Mahapatra, G., Illies, M.R., Cameron, R.A., Davidson, E.H. and  
 Etensolm, C.A.  
 A large-scale analysis of mRNAs expressed by primary mesenchyme  
 cells of the sea urchin embryo  
 Development 128 (13), 2615-2627 (2001)  
 11493577  
 JOURNAL  
 PUBMED  
 COMMENT  
 Contact: Erica Sodergren  
 Human Genome Sequencing Center  
 Baylor College of Medicine  
 One Baylor Plaza, Houston, TX 77030, USA  
 Tel: 713-798-7676  
 Fax: 713-798-6977  
 Email: erica@bcm.tmc.edu  
 NCBI Trace Archive: 506362442  
 Insert Length: 1750 Std Error: 0.25  
 Plate: 174 row: C column: 5.  
 Location/Qualifiers  
 1..857  
 /organism="Strongylocentrotus purpuratus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7668"  
 /clone="FMCSPR2-174C5"  
 /tissue\_type="embryo"  
 /cell\_type="primary mesenchyme cells"  
 /lab\_host="E.coli"  
 /clone\_lib="Sea Urchin primary mesenchyme cell cDNA  
 library"  
 /note="Vector: pSPORT1; Site\_1: NotI; Site\_2: MluI; oligo  
 dt priming from poly A+ RNA, directionally cloned"

ORIGIN

Query Match 91.1%; Score 16.4; DB 8; Length 857;  
 Best Local Similarity 94.4%; Pred. No. 3e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTTCTGGATAC 18  
 |||||  
 Db 478 CCCCCCTTTCTGGATAC 495

RESULT 24  
 AG344768/c 892 bp DNA linear GSS 18-DEC-2004  
 LOCUS Mus musculus molossinus DNA, clone: MEMG01-139F09.T7, genomic survey  
 DEFINITION  
 AG344768  
 AG344768.1 GI:47918078  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 GSS.  
 Mus musculus molossinus (Japanese wild mouse)  
 Mus musculus molossinus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T.,  
 Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriawaki, K. and  
 Shiroishi, T.  
 Contribution of Asian mouse subspecies *Mus musculus molossinus* to  
 genomic constitution of strain C57BL/6J, as defined by BAC-end  
 sequence-SNP analysis  
 Genome Res. 14 (12), 2439-2447 (2004)  
 15574823  
 JOURNAL  
 PUBMED  
 2 (bases 1 to 892)  
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
 Direct Submission  
 Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suenhiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan  
(E-mail: hatori@gscc.riken.jp, URL: http://hgp.gscc.riken.go.jp/,  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the mouse BAC library MSWg01. For BAC  
library availability, please contact Kuniya Abe (abe@cc.riken.jp).  
The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
Koyadai, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@cc.riken.jp

PRIMERS  
Sequencing : T7  
LIBRARY : PBACe3.6  
Vector : EcoRI  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI

FEATURES  
source  
1. .892  
/organism="Mus musculus molossinus"  
/mol\_type="genomic DNA"  
/sub\_species="molossinus"  
/db\_xref="taxon:57486"  
/clone="MSWg01-139F09.T7"  
/sex="male"  
/issue\_type="mixture of kidney and spleen"  
/clone\_lib="MSWg01 Mouse Male BAC Library"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 10; Length 892;  
Best Local Similarity 94.4%; Pred. No. 3e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGGATAC 18  
|||||  
46 CCCCCCTCTCTGATAC 29

Db

RESULT 25  
AG311991/c 1044 bp DNA linear GSS 18-DEC-2004  
LOCUS Mus musculus molossinus DNA, clone:MSWg01-094F09.T7, genomic survey  
DEFINITION  
AG311991  
AG311991.1 GI:47864945  
GSS.  
Mus musculus molossinus (Japanese wild mouse)  
Mus musculus molossinus  
Mus musculus molossinus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
1  
Abe, K., Noguchi, H., Tagawa, K., Yuzurika, M., Toyoda, A., Kojima, T.,  
Ezawa, K., Saitou, N., Hatori, M., Sakaki, Y., Koriwaki, K. and  
Sutroishi, T.  
Contribution of Asian mouse subspecies Mus musculus molossinus to  
genomic constitution of strain C57BL/6J, as defined by BAC-end  
sequence-SNP analysis of strain C57BL/6J, as defined by BAC-end  
Genome Res. 14 (12), 2439-2447 (2004)  
15574823  
2 (bases 1 to 1044)  
Hatori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
Direct Submission  
Submitted (17-NOV-2003) Masahira Hatori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan  
1-7-22 Suenhiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan  
(E-mail: hatori@gscc.riken.jp, URL: http://hgp.gscc.riken.go.jp/,  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the mouse BAC library MSWg01. For BAC  
library availability, please contact Kuniya Abe (abe@cc.riken.jp).  
Tsukuba Institute, Bio Resource Center,  
The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
Koyadai, Tsukuba, 305-0074 Japan

COMMENT

phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@cc.riken.jp

PRIMERS  
Sequencing : T7  
LIBRARY : PBACe3.6  
Vector : EcoRI  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI

FEATURES  
source  
1. .1044  
/organism="Mus musculus molossinus"  
/mol\_type="genomic DNA"  
/sub\_species="molossinus"  
/db\_xref="taxon:57486"  
/clone="MSWg01-094F09.T7"  
/sex="male"  
/issue\_type="mixture of kidney and spleen"  
/clone\_lib="MSWg01 Mouse Male BAC Library"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 10; Length 1044;  
Best Local Similarity 94.4%; Pred. No. 3.1e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGGATAC 18  
|||||  
119 CCCCCCTCTCTGAAAC 102

Db

RESULT 26  
BM478544 1093 bp mRNA linear EST 05-FEB-2002  
LOCUS BM478544  
AGENCOURT\_6457620 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5575228  
5', mRNA sequence.  
ACCESSION BM478544  
VERSION BM478544.1 GI:18527586  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 1093)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM12325 row: C column: 05  
High quality sequence stop: 711.  
Location/Qualifiers  
1. .1093  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5575228"  
/issue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 92"  
/note="Organ: testis; Vector: pCW-Sport6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 91.1%; Score 16.4; DB 3; Length 1093;  
Best Local Similarity 94.4%; Pred. No. 3.1e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCCCCCTCTTGATAC 18  
|||||  
811 CCCCCCTCTTGATAC 828

RESULT 27  
CK231738 1186 bp mRNA linear EST 09-DEC-2003  
LOCUS ILLUMIGEN\_MQO\_2887 Katze MMR Macaca mulatta cDNA 5' similar to  
DEFINITION human Ubiquitin Hs.417764, mRNA sequence.  
ACCESSION CK231738.1 GI:39638096  
VERSION CK231738  
KEYWORDS EST.  
SOURCE Macaca mulatta (rhesus monkey)  
ORGANISM Macaca mulatta  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Cercopithecoidea; Cercopithecinae; Macaca.  
1 (bases 1 to 1186)  
Magnes, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agay, M.B.,  
Prohl, S.C., Fitzgibbon, M., Scherer, C.A., Miner, D.G., Katze, M.G. and  
Iadonato, S.P.  
TITLE Analysis of the Macaca mulatta transcriptome and the sequence  
divergence between Macaca and human  
JOURNAL Genome Biol. 6 (7), R60 (2005)  
PUBMED 15988449  
COMMENT Contact: C. Magnes  
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA  
Tel: 2063780400  
Fax: 2063780408  
Email: cmagnes@illumigen.com  
Sequenced on 2003.11.24. 787 Q20 bases. Assembles in contig w/ 114  
member(s). Contig contains 10 (15.2%) lib members.  
PCR PRIMER  
FORWARD: CCCTCACTTAAGGGAACAA  
BACKWARD: CACTATAGCGCAATGGTA  
Insert Length: 1186 Std Error: 0.00  
Plate: CL000026 row: F column: 04  
Seq primer: CCCTCACTTAAGGGAACAA  
POLYAsites.

FEATURES  
source  
1..1186  
/organism="Macaca mulatta"  
/mol\_type="mRNA"  
/strain="Indian"  
/db\_xref="taxon:9544"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="E. coli SOLR"  
/clone\_lib="Katze MMR"  
/note="Origin: brain; Vector: Uni-ZAP XR; Site 1: Ecor I;  
Site 2: Xho I; Created from StrataGene ZAP-cDNA Synthesis  
Kit (catalog #200400) and ZAP-cDNA GigaPack III Gold  
Cloning Kit (Catalog #200450)"

ORIGIN  
Query Match 91.1%; Score 16.4; DB 7; Length 1186;  
Best Local Similarity 94.4%; Pred. No. 3.1e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCCCCCTCTTGATAC 18  
|||||  
1123 CCCCCCTCTTGGAAC 1140

Db 1123 CCCCCCTCTTGGAAC 1140

RESULT 28  
CC317382/c 1188 bp DNA linear GSS 14-MAY-2003  
LOCUS CC317382

DEFINITION TAM32-25H6 ECl.1 TAM32 Gallus gallus genomic clone TAM32-25H6,  
genomic survey sequence.  
ACCESSION CC317382  
VERSION CC317382.1 GI:30711440  
KEYWORDS GSS.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 1188)  
Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
Warren, W., Graves, T., Mardis, E. and Wilson, R.  
TITLE Gallus gallus BAC End Reads  
JOURNAL Unpublished (2003)  
COMMENT Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu  
Insert Length: 152000 Std Error: 0.00  
Seq primer: ECl TACGACTCATATAGCGG  
Class: BAC ends  
High quality sequence start: 44  
High quality sequence stop: 833.  
location/Qualifiers  
1..1188  
/organism="Gallus gallus"  
/mol\_type="genomic DNA"  
/strain="Red Jungle Fowl"  
/db\_xref="taxon:9031"  
/clone="TAM32-25H6"  
/sex="female"  
/cell\_line="UCD001, inbred 256"  
/clone\_lib="TAM32"  
/note="Vector: pECBAC1; Site 1: EcorI, Site 2: EcorI;  
TAM32 female chicken library - for library and clone  
ordering information: <http://www.hbz.tamu.edu>

ORIGIN  
Query Match 91.1%; Score 16.4; DB 9; Length 1188;  
Best Local Similarity 94.4%; Pred. No. 3.1e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCCCCCTCTTGATAC 18  
|||||  
1070 CCCCCCTCTTGATAC 1053

Db 1070 CCCCCCTCTTGATAC 1053

RESULT 29  
BM809604 1300 bp mRNA linear EST 05-MAR-2002  
LOCUS AGENCOURT\_6584097 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5467894  
DEFINITION 5', mRNA sequence.  
ACCESSION BM809604  
VERSION BM809604.1 GI:19126427  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 1300)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:



http://image.llnl.gov  
 Plate: LLCMI970 row: j column: 23  
 High quality sequence start: 37  
 High quality sequence stop: 444.  
 Location/Qualifiers

## FEATURES

source

1..1300  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5467894"  
 /issue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1b="NIH\_MGC\_41"  
 /note="Organ: skin; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAACAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 3; Length 1300;  
 Best Local Similarity 94.4%; Pred. No. 3.1e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

Db

1 CCCCCCTCTCGATAC 18  
 |||||  
 557 CCCCCCTCTCGATAC 574

## RESULT 30

BM475412

LOCUS

BM475412 1515 bp mRNA linear EST 05-FEB-2002  
 AGENCOURT\_6477789 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:5578830  
 5', mRNA sequence.

ACCESSION

BM475412

VERSION

BM475412.1 GI:18524454

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominae; Homo.

1 (bases 1 to 1515)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: CGAPBS-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNLN at:

http://image.llnl.gov

Plate: LLM12334 row: i column: 07

High quality sequence stop: 225.

Location/Qualifiers

1..1515

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5578830"

/issue\_type="retinoblastoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_1b="NIH\_MGC\_67"

/note="Organ: eye; Vector: pCMV-SPORT; Site\_1: NotI;

Site\_2: SalI; Cloned unidirectionally. Primer: oligo dT.

Average insert size 1.75 Kb. Library constructed by Life

Technologies."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 3; Length 1515;  
 Best Local Similarity 94.4%; Pred. No. 3.2e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

Db

1 CCCCCCTCTCGATAC 18  
 |||||  
 895 CCCCCCTCTCGATAC 912

## RESULT 31

CZ179143

LOCUS

DEFINITION

CZ179143 480 bp DNA linear GSS 31-JAN-2005  
 MIAA-5H14b.g1 Meloidogyne incognita BAC end sequence library  
 (MIAAGSS 001) Meloidogyne incognita genomic, genomic survey

ACCESSION

CZ179143

VERSION

CZ179143.1

KEYWORDS

GSS.

SOURCE

ORGANISM

Meloidogyne incognita (southern root-knot nematode)

Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;

Tylenchoidea; Meloidogynidae; Meloidogyninae; Meloidogyne.

1 (bases 1 to 480)

Mitreva,M., McCarter,J.P., Page,D., Martin,J., Wylie,T.,

Clifton,S., Budiman,A., Lakey,N., Opperman,C. and Bird,D.McK.

Genome Survey sequences from the parasitic nematode Meloidogyne

incognita

Unpublished (2005)

Contact: Mitreva M

Washington University in St. Louis

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: nematode@wustl.edu

BAC ends sequenced by Washington University Genome Sequencing

Center

Class: BAC ends.

Location/Qualifiers

1..480

/organism="Meloidogyne incognita"

/mol\_type="genomic DNA"

/strain="Race 1"

/db\_xref="taxon:6306"

/dev\_stage="L2"

/clone\_1b="Meloidogyne incognita BAC end sequence library

(MIAAGSS 001)"

/note="Vector: pCUGI; Site\_1: HindIII; Site\_2: HindIII;

BAC library constructed by Arlef Budiman and Nathan Lakey

at Orion Genomics, and David Bird and Charles Opperman at

Center for the Biology of Nematode Parasitism at NCSU."

ORIGIN

Query Match

Best Local Similarity

Matches

Qy

Db

3 CCCCCCTCTCGATAC 18

|||||

75 CCCCCCTCTCGATAC 90

|||||

CG585619

DEFINITION

ACCESSION

CG585619

VERSION

CG585619.1

KEYWORDS

GSS.

SOURCE

Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 536)  
Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J., Piggett, J., Beltrando-Rio, H., Buxton, E.C., Edwards, J., Finch, R.A., Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C., Key, B.W., Jr., Kipy, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D., Payter, R., Potter, D.G., Qian, N., Shaw, J., Schlick, J., Shi, Z.-Z., Sparks, M.J., Van Slichtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A.T.

TITLE Mnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

COMMENT 14610273  
Contact: Zambrowicz BP  
OmitBank

FEATURES  
source Lexicon Genetics Incorporated  
4000 Research Forest Drive, The Woodlands, TX 77381, USA  
Email: materials@lexgen.com  
Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):608-11)  
Class: Gene Trap.

ORIGIN  
Location/Qualifiers  
1..536  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129Sv/Ev"  
/db\_xref="taxon:10090"  
/clone="OST333168"  
/cell\_type="embryonic stem cell"  
/clone\_lib="Mus musculus 129Sv/Ev"

Query Match 88.9%; Score 16; DB 10; Length 536;  
Best Local Similarity 88.9%; Pred. No. 4.5e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 16 CCCCCCTTCTCGATAC 18  
16 CCCCCCTTNTNCGATAC 33

RESULT 33  
A0551912 556 bp DNA linear GSS 28-MAY-1999  
LOCUS RPCI-11-425M15.TJ RPCI-11 Homo sapiens genomic clone  
DEFINITION RPCI-11-425M15, genomic survey sequence.  
ACCESSION A0551912  
VERSION A0551912.1 GI:4911089  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 556)  
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.

AUTHORS Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building

TITLE Unpublished (1997)

JOURNAL Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org

COMMENT Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet. Co. ([info@resgen.com](mailto:info@resgen.com)). BAC end search page: [http://www.tigr.org/tcdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tcdb/hungen/bac_end_search/bac_end_search.html).  
Seq primer: SP6  
Classes: BAC clone.

FEATURES  
source Location/Qualifiers  
1..556  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7663118"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-425M15"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPCI11 Human Male BAC Library"

ORIGIN  
Query Match 88.9%; Score 16; DB 9; Length 556;  
Best Local Similarity 100.0%; Pred. No. 4.5e+03;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 CCCCCCTTCTCGATAC 18  
495 CCCCCCTTCTCGATAC 480

RESULT 34  
BP292366 582 bp mRNA linear EST 16-SEP-2004  
LOCUS BP292366 Sugano cDNA library, lung Homo sapiens cDNA clone  
DEFINITION LNC05996, mRNA sequence.  
ACCESSION BP292366  
VERSION BP292366.1 GI:52206098  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 582)  
Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Nakai, K. and Sugano, S.

AUTHORS Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions

TITLE Genome Res. 14 (9), 1711-1718 (2004)

JOURNAL Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp.

COMMENT Locution/Qualifiers

FEATURES  
source 1..582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="LNC05996"  
/issue\_type="lung"  
/clone\_lib="Sugano cDNA library, lung"

ORIGIN  
Query Match 88.9%; Score 16; DB 3; Length 582;  
Best Local Similarity 100.0%; Pred. No. 4.5e+03;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 CCCCCCTTCTCGATA 17  
143 CCCCCCTTCTCGATA 158

RESULT 35

BG027192/c 1000 bp mRNA linear EST 24-JAN-2001  
LOCUS BG027192  
DEFINITION 602295878F1 NIH\_MGC\_86 Homo sapiens cDNA clone IMAGE:4390434 5',  
mRNA sequence.  
ACCESSION BG027192  
VERSION BG027192.1 GI:12415574  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo  
1 (bases 1 to 1000)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM10079 row: 1 column: 19  
High quality sequence start: 23  
High quality sequence stop: 497.  
Location/Qualifiers  
1. 1000  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4390434"  
/issue\_type="osteosarcoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_86"  
/note="Organ: bone; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.533 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 88.9%; Score 16; DB 2; Length 1000;  
Best Local Similarity 100.0%; Pred. No. 4.8e+03;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCCCCCTTCTTGAT 16  
|||||  
Db 379 CCCCCCTTCTTGAT 364

RESULT 36  
BUI49766/c 1278 bp mRNA linear EST 03-SEP-2002  
LOCUS BUI49766  
DEFINITION AGENCOURT 8481847 NIH\_MGC\_129 Mus musculus cDNA clone IMAGE:6308007  
5', mRNA sequence.  
ACCESSION BUI49766  
VERSION BUI49766.1 GI:22663298  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1278)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Susan L. Sullivan, PhD.

CDNA Library Preparation: ResGen, Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM13726 row: 9 column: 16  
High quality sequence stop: 228.  
Location/Qualifiers  
1. 1278  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6308007"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_129"  
/note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;  
Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.  
Primer: Oligo dT. Average insert size 2.2 kb. Constructed  
by ResGen, Invitrogen Corp. Note: this is a NIH\_MGC  
Library."

ORIGIN  
Query Match 88.9%; Score 16; DB 5; Length 1278;  
Best Local Similarity 100.0%; Pred. No. 5e+03;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 CCCCCCTTCTTGATA 17  
|||||  
Db 433 CCCCCCTTCTTGATA 418

RESULT 37  
AA108376/c 73 bp mRNA linear EST 13-FEB-1997  
LOCUS AA108376  
DEFINITION mp43g05\_r1 Barethead MRLRB1 Mus musculus cDNA clone IMAGE:572024 5',  
81mler to gb: D10049 Mouse mRNA for mouse melanoma antigen,  
complete cds (MOUSE);, mRNA sequence.  
ACCESSION AA108376  
VERSION AA108376.1 GI:1659812  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 73)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [mouseest@wustl.edu](mailto:mouseest@wustl.edu)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
MGI:346672  
Seq primer: -28m13 rev2 ET from Amerham  
High quality sequence stop: 1.  
Location/Qualifiers  
1. 73  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/clone="IMAGE:572024"

FEATURES  
source



SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 192) Zhao,S., Niemann,W., Malek,J., Shatman,S., Akincet,B., Levins,M., Teegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregorgis,E., Russell,D., de Jong,P. and Fraser,C.M.
TITLE	Mouse BAC End Sequences from Library RPCI-24
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: szhaoc@igr.org Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pje@omgmai1.cno.org). Clones may be purchased from BACPAC Resources ( <a href="http://www.chori.org/bacpac/orderingframe.html">http://www.chori.org/bacpac/orderingframe.html</a> ). BAC end plate: <a href="http://www.igr.org/cdb/bac_ends/mouse/bac_end_intro.html">http://www.igr.org/cdb/bac_ends/mouse/bac_end_intro.html</a> Page: 124 row: 0 column: 17 Seq primer: SP6 Class: BAC ends.
FEATURES	Location/Qualifiers 1..192 /organism="Mus musculus" /mol_type="genomic DNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="RPCI-24-124017" /sex="Male" /cell_type="Spleen/Brain" /clone_idb="RPCI-24" /note="Vector: pGRABAC1, Site 1: BamH1, Site 2: BamH1; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pGRABAC1 cloning vector at the BamH1 sites using MboI partially digested male C57BL/6J DNA."
ORIGIN	
Query Match	85.6%; Score 15.4; DB 9; Length 192;
Best Local Similarity	94.1%; Pred. No. 7.9e+03;
Matches	16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	2 CCCCCTCTTTCGATAC 18       153 CCCCCTCTTTCGATAC 137
Db	153 CCCCCTCTTTCGATAC 137
RESULT 41	
BB160827/c	
LOCUS	BB160827 RIKEN full-length enriched, 16 days neonate thymus Mus
DEFINITION	musculus cDNA clone A130505D6 3', mRNA sequence.
ACCESSION	BB160827
VERSION	BB160827.1 GI:8816767
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Mus mycology; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 201) Komno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Iehli,Y., Iehikawa,J., Iehikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Salto,H., Sakai,C., Sato,K.,

TITLE	JOURNAL			
COMMENT	Unpublished (2000)			
	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute			
	The Institute of Physical and Chemical Research (RIKEN)			
	1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan			
	Tel: 81-45-503-9222			
	Fax: 81-45-503-9216			
	Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/			
	Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
	Thermotranscription and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)			
	Itoh, M., Kitsuina, T., Akiyama, J., Shibata, K., Itawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.			
	Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)			
	Carninci, P. and Hayashizaki, Y.			
	High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)			
	Please visit our web site (http://genome.rtc.riken.go.jp) for further details.			
FEATURES	Location/Qualifiers			
Source	1. 201			
	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/db_xref="taxon:10090"			
	/clone="A130056D06"			
	/tissue_type="thymus"			
	/dev_stage="16 days neonate"			
	/lab_host="DHI08"			
	/clone_id="RIKEN full-length enriched, 16 days neonate thymus"			
	/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' GAGGAGAGAGGATCCCAAGAGCTCTTTTCTTTTCTTNN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGGAGAGGATCTCCAGCTTAATTAATTATCCCCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a Lambda modified pBluescript KS(+) after bulk excision from Lambda			
ORIGIN	PLC I."			
	Query Match 85.6%; Score 15.4; DB 1; Length 201;			
	Best Local Similarity 94.1%; Pred. No. 7.9e+03;			
	Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0.			
QY	2 CCCCCTCTTGATAC 18			
Db	191 CCCCACCTTGTGATAC 175			
LOCUS	CKR18409			
RESULT 42	262 bp mRNA linear EST 11-MAR-2004			
CKR18409/c				

DEFINITION ic27a06.y5 HR85 islet Homo sapiens cDNA clone IMAGE: 5', mRNA  
 sequence.  
 ACCESSION CK818409  
 VERSION CK818409.1 GI:44835334  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 262)  
 REFERENCE Melton, D., Meadow, A., Clifton, S., Hillier, L., Marra, M., Page, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Treising, B., Ritzer, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V., Williams, T., Jackson, Y. and Bowers, Y.  
 WashU-Harvard Pancreas EST Project  
 TITLE Unpublished (2000)  
 JOURNAL Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 COMMENT Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@bioharp.harvard.edu  
 This read is a 5' RESEQUENCE of a previously sequenced pancreas clone  
 Good hit to opposite strand read. . wrong orientation BUT PASSED FOR MOUSE-PANCREAS VERIFICATION  
 Seq primer: -40up from Gihco  
 High quality sequence stop: 258.  
 Location/Qualifiers  
 1. 262  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:"  
 /issue\_type="Purified pancreatic islet"  
 /lab\_host="DH10B"  
 /clone\_1ib="HR85 islet"  
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1: NotI; Site\_2: XhoI; cDNA made by oligo-dT priming. XhoI site was destroyed after directional cloning. Amplified once. Contact Information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permut Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110. E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

ORIGIN  
 Query Match 85.6%; Score 15.4; DB 7; Length 262;  
 Best Local Similarity 94.1%; Pred. No. 8.2e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCCTCTCTGTGATAC 18  
 |||||  
 Db 161 CCCCTCTCTGTGATAC 145

RESULT 43  
 CE076988/c 271 bp DNA linear GSS 24-SEP-2003  
 LOCUS tigr-gss-dog-17000324369598 Dog libRARY Canis familiaris genomic,  
 DEFINITION genomic survey sequence.  
 ACCESSION CE076988  
 VERSION CE076988.1 GI:35143688  
 KEYWORDS GSS.  
 SOURCE Canis familiaris (dog)  
 ORGANISM Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;

REFERENCE 1 (bases 1 to 271)  
 AUTHORS Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K., Ruesch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and Venter, J.C.  
 TITLE The dog genome: survey sequencing and comparative analysis  
 JOURNAL Science 301 (5641), 1898-1903 (2003)  
 PUBMED 14512627  
 COMMENT Contact: Kirkness EF  
 The Institute for Genomic Research  
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-0200  
 Fax: 301-838-0208  
 Email: ekirknes@tigr.org  
 Class: Shotgun.  
 Location/Qualifiers  
 1. 271  
 /organism="Canis familiaris"  
 /mol\_type="genomic DNA"  
 /strain="Standard Poodle"  
 /db\_xref="taxon:9615"  
 /clone\_1ib="Dog Library"  
 /note="Site 1: BseXI; Libraries were prepared from peripheral blood"

FEATURES  
 source  
 1. 271  
 Location/Qualifiers

ORIGIN  
 Query Match 85.6%; Score 15.4; DB 9; Length 271;  
 Best Local Similarity 94.1%; Pred. No. 8.2e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCCTCTCTGTGATAC 18  
 |||||  
 Db 121 CCCCTCTCTGTGATAC 105

RESULT 44  
 BB399270 282 bp mRNA linear EST 15-JUL-2000  
 LOCUS BB399270 RIKEN full-length enriched, ES cells Mus musculus cDNA  
 DEFINITION Clone G330013E13 3', mRNA sequence.  
 ACCESSION BB399270  
 VERSION BB399270.1 GI:9218666  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Mus.  
 1 (bases 1 to 282)  
 REFERENCE Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurahara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toyota, T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Konno, H., et al.)  
 TITLE Unpublished (2000)  
 JOURNAL Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/



```

VERSION
KEYWORDS
BI683470..1 GI:15636413
EST.

SOURCE
ORGANISM
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS
Smith,T.P.L., Grose, W.M., Fraking, B.A., Roberts, A.J., Stone, R.T.,
Caas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C.,
Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,
Chitko-McCorm, C.G., Perera, G., Holt, I., Karaymcheva, S., Liang, F.,
Quackenbush, J. and Keel, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
11282978

JOURNAL
PUBMED

COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel.: 402 762 4366
Fax: 402 762 4390
Email: smtlhem@ars.mars.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 139 row: E column: 8
Seq primer: ATTAGTGACACTTATG.
Location/Qualifiers
1..306
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_id="MARC 1BOV"
/notes="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

ORIGIN
Query Match 85.6%; Score 15.4; DB 3; Length 306;
Bos Local Similarity 94.1%; Pred. No. 8.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCTCTTCGGATAC 18
|||||
Db 78 CCCCTCTTCGGATGC 62

RESULT 47
LOCUS AM883754 313 bp mRNA linear EST 23-MAY-2000
DEFINITION QV-OTD062-090500-207-E03 OT0062 Homo sapiens CDNA, mRNA sequence.
ACCESSION AM883754
VERSION AM883754.1 GI:8045766
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 313)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Birones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed

```

JOURNAL	sequence tags
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
COMMENT	10737800 Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-27049922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the PAPER/PtCR Human Cancer Genome Project. This entry can be seen in the following URL <a href="http://www.ludwig.org.br/scripts/gethtml.pl?file=QV2-QT0062-090">http://www.ludwig.org.br/scripts/gethtml.pl?file=QV2-QT0062-090</a> 500-207-03k&t3=2000-05-09&c4=1) Seq primer: puc 18 forward High quality sequence start: 96 High quality sequence stop: 311. Location/Qualifiers
FEATURES	1..313
SOURCE	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /dev_stage="Adult" /clone_1lb="OT0062" /notes="Organ: ovary; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
ORIGIN	
Query Match	85.6%; Score 15.4; DB 1; Length 313;
Best Local Similarity	94.1%; Pred. No. 8.4e+03;
Matches	16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy	2 CCCCCCTTTCTGCATAC 18       CCCCCTTTCTGGAGAC 86
Dn	102 CCCCCCTTTCTGGAGAC 86
RESULT 48	
LOCUS	BB224660 325 bp mRNA linear EST 01-JUL-2000
DEFINITION	BB224660 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA AS30087J33 3', mRNA sequence.
ACCESSION	BB224660
VERSION	BB224660.1 GI:8693272
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus. 1 (bases 1 to 325) Kono,H., Alzawa,K., Akahira,S., Akiyama,T., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hiroyane,T., Ishii,F., Ishii,Y., Ishikawa,D., Ishikawa,T., Itoh,M., Itawa,M., Kadori,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Oda,H., Kusakaabe,M., Matsuyama,T., Miki,R., Mituno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tsegawa,A., Takahashi,F., Tomimaga,N., Toyota,T., Tsunoda,Y., Watanahi,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasuniishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	RIKEN Mouse ESTs (Kono,H., et al.)
JOURNAL	Unpublished (2000)
COMMENT	Contact: Yoshihide Hayashizaki





Best Local Similarity 94.1%; Pred. No. 8.4e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCTCTCTGATAC 18  
104 CCCCTCTCTGATAC 120

RESULT 50  
A2722361/c 328 bp DNA linear GSS 24-JAN-2001

LOCUS RPCI-24-9319.TJ RPCI-24 Mus musculus genomic clone RPCI-24-9319,  
DEFINITION genomic survey sequence.

ACCESSION A2722361  
VERSION A2722361.1 GI:12465964  
KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Murioidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 328)  
Zhao,S., Niernan,W., Malek,J., Shatsman,S., Akintet,B., Levins,M.,  
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgie,E.,  
Russell,D., de Jong,P. and Fraser,C.M.

AUTHORS Mouse BAC End Sequences from Library RPCI-24

TITLE Unpublished (1999)

JOURNAL Other GSSs: RPCI-24-9319.TV

COMMENT Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208

Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC  
library availability, please contact Pieter de Jong  
(pdejong@mail.cho.org). Clones may be purchased from BACPAC  
Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end  
plates: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
Plate: 93 row: 1 column: 9  
Seq primer: SPE  
Class: BAC ends.

FEATURES  
source Location/Qualifiers

1..328  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-9319"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/clone\_lib="RPCI-24"  
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI;  
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The  
library was cloned in the pTARBAC1 cloning vector at the  
BamHI sites using MboI partially digested male C57BL/6J  
DNA."

ORIGIN

Query Match 85.6%; Score 15.4; DB 9; Length 328;  
Best Local Similarity 94.1%; Pred. No. 8.4e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCTCTCTGATAC 17  
196 CCCCTCTCTGATAC 180

RESULT 51

LOCUS AW482062/c 329 bp mRNA linear EST 25-APR-2001  
DEFINITION 40528 MARC 3BOV Bos taurus CDNA 5', mRNA sequence.

ACCESSION AW482062  
VERSION AW482062.1 GI:7052168  
KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 329)  
Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Caasas,E., Wray,J.B., White,J., Cho,J., Fahrenkrug,S.C.,  
Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A.,  
Chicko-McKown,C.G., Pettea,G., Holt,I., Katamycheva,S., Liang,F.,  
Quackenbush,J. and Keefe,J.W.

AUTHORS Sequence evaluation of four pooled-tissue normalized bovine CDNA  
libraries and construction of a gene index for cattle

TITLE Genome Res. 11 (4), 626-630 (2001)

JOURNAL 11282978

COMMENT Contact: Smith TPJ  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390

Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 20  
and -mismatch 12 options.

PCR Primers  
FORWARD: AGGAAACAGCTATGACCAT  
BACKWARD: GTTTCACGTCACGACG  
Plate: 20 row: C column: 5  
Seq primer: ATTGAGTGCACCTATAG.

FEATURES  
source Location/Qualifiers

1..329  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="MARC 3BOV"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
library made from pooled tissue from marrow, alveolar  
macrophage, ovary, fetal semitendinous muscle, and fecal  
longissimus muscle."

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 329;  
Best Local Similarity 94.1%; Pred. No. 8.4e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCTCTCTGATAC 18  
255 CCCCTCTCTGATAC 239

RESULT 52  
LOCUS BU185054/c 352 bp mRNA linear EST 16-OCT-2003

DEFINITION BU185054 normalized full length cDNA library, chloromimata,  
caulonemata and malformed buds Physcomitrella patens subsp. patens  
CDNA clone pphb5j19 5', mRNA sequence.

ACCESSION BU185054  
VERSION BU185054.1 GI:18352999

KEYWORDS EST.

ORGANISM Physcomitrella patens subsp. patens  
Physcomitrella patens subsp. patens  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.

REFERENCE 1 (bases 1 to 352)  
Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H.,  
Uchiyama,T., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,  
Kohara,Y. and Hasebe,M.

AUTHORS Comparative genomes of Physcomitrella patens gametophytic

JOURNAL  
PUBMED  
COMMENT

transcriptome and Arabidopsis thaliana: implication for land plant evolution  
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)  
12808149  
Contact: Tadasu Shin-i  
Center for Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp  
A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carlini et al., 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'-gagagagagagatccacccctggagagatgtttttttttttttt-3' was used as a 1st 3' primer, and 5'-gggttcgagatcgatcgctgttccagacagcagatgactcgagAACGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAATCGCCGCGATCGAATCGCGAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).  
Protonemata were blended by the POLYTRON, and then cultivated on the BODDAG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PhysCObase (<http://mos.nibb.ac.jp>).  
Location/Qualifiers

FEATURES  
source

```
1..352
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/db_xref="taxon:145481"
/clone="pPhb5j19"
/tissue_type="mixture of chloronemata, caulonemata and malformed buds"
/clone_lib="normalized full length cDNA library, chloronemata, caulonemata and malformed buds"
```

## ORIGIN

Query Match 85.6%; Score 15.4; DB 3; Length 352;  
Best Local Similarity 94.1%; Pred. No. 8.5e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCTCTTCTGGATAC 18  
|||  
Db 239 CCCCTCTTCTGGATCC 223

RESULT 53  
C0878314/c 356 bp mRNA linear EST 01-SEP-2004  
LOCUS  
DEFINITION  
RZPDp1056122290 5', mRNA sequence.

ACCESSION  
C0878314  
VERSION  
C0878314.1 GI:51808234  
KEYWORDS  
EST.  
SOURCE  
Bos taurus (cow)  
ORGANISM  
Bos taurus

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
Hennig S., Janitz M., Herwig R. and Williams J.  
Generation, annotation, evolutionary analysis and database integration of 14969 cattle EST clusters  
Unpublished (2004)  
Contact: Hennig S  
laboratory 123, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik

Inhestr 63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1612  
Fax: +49 30 8413 1380  
Email: hennig@olgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones that display the same hybridisation matrix with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per ONFP cluster was selected for sequencing. cDNA clones and filters are distributed via Deutsches Ressourcenzentrum fuer Genomforschung GmbH (<http://www.rzpd.de>).

PCR Primers  
FORWARD: 5' CCCGAGCTTACCTTATGCTCCGGCTCG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GGTATTAACCCGCTGGCGAAGGGGAGATG 3' (M13FSP) 3'-seq  
Seq primer: 5'-CCGCTCGGATTCGGGGT-3' (M13RSP).

FEATURES  
source

```
1..356
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="RZPDp1056122290"
/sex="female"
/tissue_type="brain tissue"
/dev_stage="adult brain"
/clone_lib="normal cattle brain"
/clone="Organ: Brain; Vector: pSport1; Site 1: NotI; Site 2: SalI; Random primed and directionally cloned in pSport1 vector using NotI (5'-pGACTGATCTGATCGAGCGGCGGCC (T)15-3' and SalI 5'-TCGACCCACCGCTCCG-3' adaptors (Gibco BRL)."
```

## ORIGIN

Query Match 85.6%; Score 15.4; DB 7; Length 356;  
Best Local Similarity 94.1%; Pred. No. 8.5e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCTCTTCTGGATAC 18  
|||  
Db 109 CCCCTCTTCTGGATGC 93

RESULT 54  
CN439166/c 379 bp mRNA linear EST 07-JUN-2004  
LOCUS  
DEFINITION  
BE04018A2B01 BE04 Normalized and Subtracted Bovine Embryonic and extraembryonic tissue Bos taurus cDNA clone BE04018A2B01 5', mRNA sequence.

ACCESSION  
CN439166  
VERSION  
CN439166.1 GI:46418430  
KEYWORDS  
EST.  
SOURCE  
Bos taurus (cow)  
ORGANISM  
Bos taurus

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
1 (bases 1 to 379)  
Lewin H.A., Renard J.P., Yang X.J., Hernandez A., Degrelle S., Hue I., Tian X.C., Liu L. and Everts R.E.  
Bovine embryonic ESTs  
Unpublished (2004)  
Contact: Harris Lewin  
Department of Animal Sciences  
University of Illinois at Urbana-Champaign  
206 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL 61801, USA  
Tel: 217 333 5998  
Fax: 217 244 5617  
Email: h-lewin@uiuc.edu  
Funding for embryonic EST sequencing was provided by USDA ARS contract No. 58-1275-2-020 to H. A. Lewin. Base Calling/Quality Scores: PHRED from Washington University Genome Center  
Trimming: Cross\_match from Washington University Genome Center  
PhRAP suite. Sequences submitted are vector free and at least 200

bp in length with average PHRED score > 20.  
 PCR Primers  
 FORWARD: CCAGTCAGACCTTGTAAACGAC (M13 NTA-F)  
 BACKWARD: GTGTGAATGTGACGATACAA (M13 NTA-R)  
 Insert length: 379 Std Error: 0.00  
 Plate: BE04018A2 row: B column: 01  
 Seq primer: TATACGACTACTATAGG (77 PROMOTER)  
 High quality sequence stop: 379.  
 Location/Qualifiers

## FEATURES

source

1..379  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /clone="BE04018A2B01"  
 /tissue\_type="embryo (day 36 and day 64) and extra-embryonic tissue (day 14 to 25) Normalized and twice subtracted library"  
 /lab\_host="DH10B"  
 /clone\_lib="BE04 Normalized and Subtracted bovine embryonic and extraembryonic tissue"  
 /notes="vector: pGEMZf1(+); Site\_1: EcoRI; Site\_2: NotI; The cDNA library was constructed, normalized and subtracted by Dr. A. Hernandez, W.M. Keck Center, University of Illinois Urbana-Champaign and S. Degrelle, BPR, Institut National de la Recherche Agronomique, France, as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. This clone is from library BE04, a normalized library of embryonic tissues subtracted with 7800 known placental EST sequences and 5000 clones from the BE03 library. The double stranded cDNA was size selected (more than 450 bp), adapted with EcoRI adaptors at both ends and then digested with NotI. The cDNA was then directionally cloned into EcoRI-NotI digested pGEMZf1(+). phagendic vector. Insert size was between 450-2200 bp. Sequencing was done from the 5' end of the clone. Tagged oligo-dT was used to identify the source of the ESTs where possible. (A18)TGGCT = extraembryonic tissue; (A18)TGGCT = 36-day old fetus; (A18)TGGCA = 64-day old fetus (organs); (A18)TGGCA = 64-day old fetus (body)"

## ORIGIN

Query Match 85.6%; Score 15.4; DB 6; Length 381;  
 Best Local Similarity 94.1%; Pred. No. 8.6e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCCTCTTCTGATAC 18  
 |||||  
 Db 125 CCCCCTCTTCTGATGC 109

RESULT 55  
 LOCUS CB423611 381 bp mRNA linear EST 25-MAR-2003  
 DEFINITION 597072 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.  
 ACCESSION CB423611  
 VERSION CB423611.1 GI:29192827  
 KEYWORDS EST.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
 Pecora; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 381)  
 Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,  
 Wray,J.E. and Keele,J.W.

TITLE A second set of bovine ESTs from pooled-tissue normalized libraries  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390

Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called with phred v0.020425.c and  
 trimmed with the aid of the trim\_alt option. Vector identified with  
 cross\_match v0.990329.  
 Plate: FQ18005 row: B column: 15  
 Seq primer: TAGAAGCACAGTCGAGG.  
 Location/Qualifiers

## FEATURES

source

1..381  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /clone\_lib="MARC 6BOV"  
 /note="vector: pCDNA3.1; Site\_1: EcoRI; Site\_2: NotI; library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."

## ORIGIN

Query Match 85.6%; Score 15.4; DB 6; Length 381;  
 Best Local Similarity 94.1%; Pred. No. 8.6e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCCTCTTCTGATAC 18  
 |||||  
 Db 254 CCCCCTCTTCTGATGC 270

RESULT 56  
 LOCUS AF227818 384 bp mRNA linear HTC 23-JAN-2003  
 DEFINITION Ictalurus punctatus clone CM191 mRNA, partial sequence.  
 ACCESSION AF227818  
 VERSION AF227818.1 GI:27883583  
 KEYWORDS HTC.  
 SOURCE Ictalurus punctatus (channel catfish)  
 ORGANISM Ictalurus punctatus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;  
 Ictaluridae; Ictalurus.

REFERENCE 1 (bases 1 to 384)  
 Kim,S., Li,P., Zheng,X., Dunham,R.A. and Liu,Z.  
 Gene expression in the muscles of young and mature channel catfish  
 (Ictalurus punctatus) as analyzed by expressed sequence tags and  
 gene filters

JOURNAL Fish Physiol. Biochem. (2003) In press  
 REFERENCE 2 (bases 1 to 384)  
 Liu,Z. and Kim,S.

TITLE Direct Submission  
 JOURNAL Submitted (21-JAN-2000) Fisheries and Allied Aquacultures and the  
 Program of Cell and Molecular BioSciences, Auburn University, 203  
 Swingle Hall, Auburn, AL 36849, USA

FEATURES  
 source 1..384  
 /organism="Ictalurus punctatus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7998"  
 /clone="CM191"  
 /tissue\_type="muscle"

## ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 384;  
 Best Local Similarity 94.1%; Pred. No. 8.6e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCTCTTCTGATA 17  
 |||||  
 Db 129 CCCCCTCTTCTGAGA 145

RESULT 57  
 BFs44009/c

LOCUS	BF544009	386 bp	mRNA	linear	EST 11-DEC-2000
DEFINITION	UI-R-E0-Cg-f-10-0-U1.r1 UI-R-E0 Rattus norvegicus cDNA clone				
ACCESSION	UI-R-E0-Cg-f-10-0-U1.5				mRNA sequence.
VERSION	BF544009				
KEYWORDS	BF544009.1 GI:11635116				EST.
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus. 1 (bases 1 to 386) Bonaldo,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9), 791-806 (1996)				
AUTHORS	Contact: Soares, MB				
JOURNAL	Coordinated Laboratory for Computational Genomics				
PUBMED	University of Iowa 375 Newton Road , 4156 MBRRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu				
COMMENT	cDNA library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.regen.com) This clone is also available through the I.M.A.G.E.B. Consortium at LMU (inf@image.lmu.gov). IMAGE ID=1791950 The following repetitive elements were found in this cDNA sequence: 13-67, >PBD9SINE/Alu 14-134, >Bl-FSINE/Alu 149-220, >ID2ASINE/ID 278-380, >Bl_MM5SINE/Alu Seq primer: M13 Forward Location/Qualifiers 1..386 /organism="Rattus norvegicus" /mol_type="mRNA" /strain="Sprague-Dawley" /db_xref="taxon:10116" /clone="UI-R-E0-Cg-f-10-0-U1" /dev_stage="embryonic" /lab_host="DH10B (Life Technologies)" /clone_lib="UI-R-E0" /note="Vector: pRT3D-Pac (pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; This library consists of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture."				
ORIGIN					
Query Match	85.6%;	Score 15.4;	DB 2;	Length 386;	
Best Local Similarity	94.1%;	Ped. No. 8.6e+03;			
Matches 16;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0.	
CY	2 CCCCTCTTCGTGATAC 18				
Dd					
	326 CCCCACCTTCGGATAC 310				
RESULT 58					
LOCUS	BM734793	386 bp	mRNA	linear	EST 01-MAR-2000
DEFINITION	MONOI_13_H02_g1_A005 Monocytes (MONOI) Equus caballus cDNA, mRNA sequence.				
ACCESSION	BM734793				
VERSION	BM734793.1 GI:19056126				
KEYWORDS	EST.				
SOURCE	Equus caballus (horse)				
ORGANISM	Equus caballus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus				

REFERENCE	1 (bases 1 to 386)					
AUTHORS	Vanderplas M.L., Cordonnier-Pratt,M.-M., Sudman,M.L., Wentzel,V.E., Gingle,A.R., Pratt,L.H. and Moore,J.N.					
TITLE	An EST database from equine ( <i>Equus caballus</i> ) monocytes unpublished (2001)					
JOURNAL	Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210					
COMMENT	Email: mmprratt@ga.edu Sequences have been trimmed to exclude polyA, vector and regions below Phred quality 16. The threshold for high quality sequence is 20. Three-prime sequences, which are obtained with PolyTmix or T7 sequencing primer, are presented as the reverse complement. Seq primer: T7 High quality sequence stop: 386 POLYA=yes.					
FEATURES						
SOURCE	location/Qualifiers					
	1..386					
	/organism="Equus caballus"					
	/mol_type="mRNA"					
	/db_xref="taxon:9796"					
	/cell_type="Isolated peripheral blood monocytes stimulated with E. coli lipopolysaccharide"					
	/clone_lib="Monocytes (MONO1)"					
	Note="Vector: p Bluescript SK(-) from Lambda ZapII; site_1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda ZapII. Clones to be sequenced were prepared by mass excision."					
ORIGIN						
	Query Match	85.6%;	Score 15.4;	DB 3;	Length 386;	
	Best Local Similarity	94.1%;	Pred. No. 8.6e+03;			
	Matches 16;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
Oy	1 CCCCCCTTCGTGATA 17					
Dd	6 CCCCCTCCTTGAGAGA 22					
RESULT 59						
CK443748	387 bp mRNA linear EST 29-JUN-2005					
LOCUS	GQ0015c.BB.1.D11.GQ001: Male strobili developmental sequence Picea					
DEFINITION	glauca cDNA clone GenomeQuebec_Id:GQ0015CD11 3', mRNA sequence.					
ACCESSION	CK443748					
VERSION	CK443748.1 GI:40780168					
KEYWORDS	EST.					
SOURCE	Picea glauca (white spruce)					
ORGANISM	Picea glauca					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Coniferopsida; Coniferales; Pinaceae; Picea.					
AUTHORS	1 (bases 1 to 387) Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A., Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stoltz,J., Sidiqi,A., Holt,R., Marra,M. and Mackay,J.					
TITLE	Arbores EST sequencing in Picea glauca (white spruce) unpublished (2004)					
JOURNAL	Contact: John Mackay					
COMMENT	Centre de Recherche en Biologie Forestiere Universite Laval Pavillon Charles-Eugene Marchand, Quebec, CANADA G1K 7P4 Fax: 418 656 7493 Email: jmackay@rsvs.ulaval.ca Center for Computational Genomics and Bioinformatics (CCGB), University of Minnesota, MN_id Identifier: MN5160261 Clone ID: GQ0015c D11 Clones available through: John Mackay, Ph. D. Professeur adjoit -Assistant professeur EMAIL: jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere (Forest Biology Research center) Universite Laval Quebec, Quebec CANADA G1K 7P4					

Plate: 5c row: 11 column: D  
Seq primer: PolyTplus primer.  
Location/Qualifiers

FEATURES  
source  
1..387  
/organism="Picea glauca"  
/mol\_type="mRNA"  
/strain="Tree 13-271"  
/db\_xref="taxon:3330"  
/clone="GenomeQuebec\_Id:GQ0015cd11"  
/sex="Hermaphrodite"  
/tissue\_type="Entire strobilus"  
/dev\_stage="Three stages of preformed male strobili at end of winter dormancy were pooled: swollen fully closed buds, partly open buds and fully open buds"  
/lab\_host="E. coli DH10B cells"  
/clone\_lib="GQ001: Male strobili developmental sequence"  
/note="Organ: Expanding male strobili; Vector: plusscript II SK (+) XR; Site 1: Eco-RI; Site 2: Xho-I; cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the Bluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitro) for propagation"

## ORIGIN

Query Match 85.6%; Score 15.4; DB 7; Length 387;  
Best Local Similarity 94.1%; Pred. No. 8.6e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCCTCTTCGTGATAC 18  
|||||  
Db 300 CCCCCTCTTCGTGATAC 316

RESULT 60  
CE777815 390 bp DNA linear GSS 30-SEP-2003  
LOCUS CE777815  
DEFINITION tigr-gss-dog-1700031697596 Dog library Canis familiaris genomic,  
genomic survey sequence.

ACCESSION CE777815  
VERSION CE777815.1 GI:37118578  
KEYWORDS GSS.  
SOURCE Canis familiaris (dog)  
ORGANISM Canis familiaris

REFERENCE  
AUTHORS 1 (bases 1 to 390)  
Kirkenes, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K.,  
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
Venter, J.C.  
TITLE The dog genome: survey sequencing and comparative analysis  
JOURNAL Science 301 (5641), 1898-1903 (2003)  
PUBMED 14512627  
COMMENT Contact: Kirkenes EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirkenes@tigr.org  
Class: shotgun.

FEATURES  
source  
Location/Qualifiers

1..390  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

## ORIGIN

Query Match 85.6%; Score 15.4; DB 10; Length 390;

Best Local Similarity 94.1%; Pred. No. 8.6e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCCTCTTCGTGATAC 18  
|||||  
Db 85 CCCCCTCTTCGTGATAC 101

## RESULT 61

BY472844 398 bp mRNA linear EST 02-JAN-2003  
LOCUS BY472844  
DEFINITION BY472844 RIKEN full-length enriched, melanocyte mus musculus cDNA  
clone G270125H13 3', mRNA sequence.

ACCESSION BY472844  
KEYWORDS BY472844.1 GI:26807223  
SOURCE EST.  
MUS musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridea; Muridae; Murinae; Mus.

## REFERENCE

## AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,  
Chochia, C., Corbani, L.E., Cousine, S., Dalla, E., Dragani, T.A.,  
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gariboldi, M., Glasi, C., Godzik, A., Gough, J., Grifmond, S.,  
Gustincin, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,  
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Semple, C.A., Secou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
Hayata, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, T., Miyazaki, A., Sasaki, K., Sasaki, D., Shibata, K.,  
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
Rogers, J., Birney, E. and Hayashizaki, Y.  
TITLE Analysis of the mouse transcriptome based on functional annotation  
JOURNAL Nature 420, 563-573 (2002)  
PUBMED 12466851  
COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,  
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,  
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
Shiraki, T., Tagami, M., Waki, K., Watanabe, M., Watanabe, M. and  
Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Cells were provided by Drs. William J Pavan, Stacie Loftus, and Denise Larson (Division of Intramural Research Genetic Disease Research Branch National Human Genome Research Institute, National Institutes of Health (NIH) Building: 49, Room 4A82 49 Convent Drive MSC 4472 Bethesda, Maryland U.S.A) whose assistance we gratefully acknowledge.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

## source

Location/Qualifiers  
 1..398  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="G270125H13"  
 /cell\_type="melanocyte"  
 /clone\_1lb="RIKEN full-length enriched, melanocyte"

## ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 398;  
 Best Local Similarity 94.1%; Pred. No. 8.6e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 CCCCCTTCTCGATAC 18  
 |||||  
 169 CCCCCTTCTCGACAC 185

Db 169 CCCCCTTCTCGACAC 185

## RESULT 62

## BY702625

LOCUS BY702625 RIKEN full-length enriched, whole joints Mus musculus cDNA

DEFINITION clone L230004K16 3', mRNA sequence.

ACCESSION BY702625

VERSION BY702625.1 GI:27113718

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

399 bp mRNA linear EST 16-DEC-2002  
 BY702625 RIKEN full-length enriched, whole joints Mus musculus cDNA  
 clone L230004K16 3', mRNA sequence.  
 BY702625  
 BY702625.1 GI:27113718  
 EST.  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
 1 (bases 1 to 399)  
 Okazaki, T., Furuno, M., Kaenkawa, T., Adachi, J., Bono, H., Kondo, S., Nakada, H., Ota, N., Saito, R., Suzuki, H., Yamana, T., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schiraldi, L. M., Kanihara, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusic, V., Chochina, C., Corbani, L. E., Cousins, S., Dail, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Guenichon, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawasaki, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Kongaya, A., Kurachkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltara, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numa, K., Okido, T., Pavan, W. J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. V., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Savelle, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, J., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wysshaw-Boris, A., Yangisawa, M., Yang, T., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

TITLE  
 JOURNAL  
 PUBMED  
 COMMENT  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: [genome-res@sc.riken.jp](mailto:genome-res@sc.riken.jp), [URL:http://genome.gsc.riken.jp/](mailto:URL:http://genome.gsc.riken.jp/)  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hayazaki, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, Y., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shizaki, T., Tagami, M., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Tissues were provided by Vassilis Aidinis (Biomedical Sciences Research Center 'A1. Fleming' Institute of Immunology 14-16 A1, Fleming street 16672 Vari, Greece) whose assistance we gratefully acknowledge.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

## source

Location/Qualifiers  
 1..399  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="L230004K16"  
 /tissue\_type="whole joints"  
 /clone\_1lb="RIKEN full-length enriched, whole joints"

## ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 399;  
 Best Local Similarity 94.1%; Pred. No. 8.6e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 CCCCCTTCTCGATAC 18  
 |||||  
 108 CCCCCTTCTCGATAC 124

Db 108 CCCCCTTCTCGATAC 124

## RESULT 63

## BM124901/c

LOCUS BM124901 400 bp mRNA linear EST 01-FEB-2002  
 DEFINITION L0546H05-3 NIA Mouse Newborn Heart cDNA Library Mus musculus cDNA  
 clone L0546H05 3', mRNA sequence.

ACCESSION BM124901

VERSION BM124901.1 GI:17108669

## KEYWORDS

EST.

SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Pantano, S., Lim, M.K. and Ko, M.S.H.  
TITLE Systematic Analyses of NIA Mouse Newborn Heart cDNA Library  
JOURNAL Unpublished (2001)  
COMMENT Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igeni.gsc.nia.nih.gov  
Plate: L0546 row: H column: 05  
Seq primer: -21M13 Forward  
High quality sequence stop: 400  
POLYA=Yes.

FEATURES  
source  
Location/Qualifiers  
1..400  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/db\_xref="taxon:10090"  
/clone="L0546H05"  
/issue\_type="Newborn Heart"  
/dev\_stage="Newborn"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse Newborn Heart cDNA Library"  
/note="Organ: heart; Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igeni.gsc.nia.nih.gov/cDNA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-  
pACTGATGTTAGATCGCAGCGCGCCCTTTTCTTTT-3'] from 24.9 microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L0-Sal3 (Ref. Development 127:1737-1749 (2000) [PMID: 10725249]), purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were digested with SalI and NotI enzymes, and cloned into SalI and NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with ligation mixture by the chemical method. The average insert size is about 1.8 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN  
Query Match 85.6%; Score 15.4; DB 3; Length 400;  
Best Local Similarity 94.1%; Pred. No. 8.6e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTTCGATA 17  
|||||  
Db 162 CCCCCCTCTTTCGATA 146

RESULT 64  
CK443747/c 405 bp mRNA linear EST 29-JUN-2005  
LOCUS GQ0015C.BR.D11.GQ001: Male strobili developmental sequence Picea  
DEFINITION glauca cDNA clone GenomeQuebec\_Id:GQ0015CD11 5', mRNA sequence.  
ACCESSION CK443747  
VERSION CK443747.1 GI:40780167  
KEYWORDS EST.  
SOURCE Picea glauca (white spruce)  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.  
REFERENCE 1 (bases 1 to 405)

AUTHORS Morency, M.-J., Cooke, J., Pavy, N., Parsons, L., Paule, C., Seguin, A., Rezel, E., Butterfield, Y., Barber, S., Yang, G., Stott, J., Siddiqui, A., Holt, R., Mair, M. and Mackay, J.  
TITLE Arborea EST sequencing in Picea glauca (white spruce)  
JOURNAL Unpublished (2004)  
COMMENT Contact: John Mackay  
Centre de Recherche en Biologie Forestiere  
Universite Laval  
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
Fax: 418 656 7493  
Email: jmackay@revs.ulaval.ca  
Center for Computational Genomics and Bioinformatics (CCGB),  
University of Minnesota, MN id Identifier: MN5160260 Clone ID:  
GQ0015C.D11 Clones available through: John Mackay, Ph. D.  
Professeur adjoint - Assistant professeur EMAIL:  
jmackay@revs.ulaval.ca Centre de Recherche en Biologie Forestiere  
(Forest Biology Research Center) Universite Laval Quebec, Quebec  
CANADA G1K 7P4  
Plate: 5c row: 11 column: D  
Seq primer: M13 Reverse Primer.

FEATURES  
source  
Location/Qualifiers  
1..405  
/organism="Picea glauca"  
/mol\_type="mRNA"  
/strain="Tree 13-271"  
/db\_xref="taxon:3330"  
/clone="GenomeQuebec\_Id:GQ0015CD11"  
/sex="Hermaprodite"  
/issue\_type="Entire strobilus"  
/dev\_stage="Three stages of preformed male strobili at end of winter dormancy were pooled: swollen fully closed buds, partly open buds and fully open buds"  
/lab\_host="E. coli DH10B cells"  
/clone\_lib="GQ001: Male strobili developmental sequence"  
/note="Organ: Expanding male strobili; Vector: pBluescript II SK (+) XR; Site 1: Eco-RI; Site 2: Xho-I; cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitro) for propagation"

ORIGIN  
Query Match 85.6%; Score 15.4; DB 7; Length 405;  
Best Local Similarity 94.1%; Pred. No. 8.6e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCCCTCTTTCGATAC 18  
|||||  
Db 79 CCCCCCTCTTCGATAC 63

RESULT 65  
BJ197399/c 414 bp mRNA linear EST 21-OCT-2003  
LOCUS BJ197399  
DEFINITION BJ197399 normalized full length cDNA library, chloronemata, caulemema and rhizoid-like pronemema Physcomitrella patens subsp. patens cDNA clone pphn29013 5', mRNA sequence.  
ACCESSION BJ197399  
VERSION BJ197399.1 GI:18365325  
KEYWORDS EST.  
SOURCE Physcomitrella patens subsp. patens  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.  
REFERENCE 1 (bases 1 to 414)  
AUTHORS Nishiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe, M.  
TITLE Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)  
PubMed 12808149





amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003). Protonemata were blended by the POLYTRON, and then cultivated on the BCGATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCObase (<http://moss.nibb.ac.jp/>).

## FEATURES

source

1. .416  
/organism="Physcomitrella patens subsp. patens"  
/mol\_type="mRNA"  
/sub\_species="patens"  
/db\_xref="taxon:145481"  
/clone="pPhn42k20"  
/tissue\_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"  
/clone\_lib="normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata"

## ORIGIN

Query Match 85.6%; Score 15.4; DB 3; Length 416;  
Best Local Similarity 94.1%; Pred. No. 8.7e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## QY

2 CCCCTCTTCTGGATAC 18  
|||||  
239 CCCCTCTTCTGGATCC 223

## DB

RESULT 68  
BU917350/c 418 bp mRNA linear EST 30-MAR-2004

## LOCUS

EST292 Bovine Lambda Zap Express corpus luteum cDNA library Bos taurus cDNA clone clt\_006\_h09\_5' similar to H. sapiens mRNA for Gal-beta (1-3/1-4)GlcNAc alpha-2,3-sialyltransferase, mRNA sequence.

## ACCESSION

BU917350  
BU917350.1 GI:45825734

## VERSION

EST.

## KEYWORDS

Bos taurus (cow)

## SOURCE

Bos taurus

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

## REFERENCE

AUTHORS

Casey, O.M., Fitzpatrick, R., McInerney, J.O., Morris, D.G., Powell, R. and Sreenan, J.M.

## TITLE

Analysis of gene expression in the bovine corpus luteum through generation and characterisation of 960 ESTs

## JOURNAL

PUBMED

15245912

## COMMENT

Contact: Casey OM  
Animal Reproduction Department  
Teagasc, Agriculture and Food Development Authority  
Galway, Ireland  
Tel: 00353 91 845845  
Fax: 00353 91 845847  
Email: ocasey@athenry.teagasc.ie  
Insert Length: 418 Std Error: 0.00  
Plate: corpus luteum rate plate 6 row: h column: 09  
Seq primer: M13 reverse primer = caggaacagctatgacc  
High quality sequence stop: 418  
POLYA=yes.

## FEATURES

source

Location/Qualifiers  
1. .418  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/clone="clt\_006\_h09"  
/sex="Female"  
/tissue\_type="Corpus luteum"

/dev stage="Day 6, 8 and 14 of the oestrus cycle"  
/clone\_lib="Bovine Lambda Zap Express corpus luteum cDNA library"  
/note="Organ: Corpus luteum; Vector: Lambda ZAP Express; Site 1: EcoRI, Site 2: XhoI; A bovine corpus luteum cDNA library was constructed using the Lambda Zap Express/GigaPack cloning kit (Stratagene cloning systems). cDNA synthesis was carried out using an oligo-(dT) primer for the reverse transcription of 5'g of mRNA and the library was constructed by directional cloning EcoRI-XhoI based on manufacturers instructions. An insert:vector ligation ratio of 1:5 was chosen as most optimum. The lambda library was packaged with GigaPack III gold packaging extracts and plated on the E. coli cell line XL1-Blue MRF".

## ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 418;  
Best Local Similarity 94.1%; Pred. No. 8.7e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## QY

2 CCCCTCTTCTGGATAC 18  
|||||  
172 CCCCTCTTCTGGATGC 156

## DB

## RESULT 69

AV667682

## LOCUS

AV667682 Bos taurus ovary fetus Bos taurus cDNA clone E10V017B05 3', mRNA sequence.

## ACCESSION

AV667682

## VERSION

AV667682.1 GI:9932427

## KEYWORDS

Bos taurus (cow)

## SOURCE

Bos taurus

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

## REFERENCE

AUTHORS

Takahara, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H. and Sugimoto, Y.

## TITLE

Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs

## JOURNAL

PUBMED

11713328

## COMMENT

Contact: Yoshikazu Sugimoto  
Animal Genetics Division  
Shirakawa Institute of Animal Genetics  
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan  
Tel: 81-248-25-5641  
Fax: 81-248-25-5725  
Email: kazueugi@siag.or.jp  
Single pass sequencing.  
This clone was obtained from a polyA-deleted cDNA library.

## FEATURES

source

Location/Qualifiers  
1. .419  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/clone="E10V017B05"  
/tissue\_type="ovary"  
/dev\_stage="fetus"  
/lab\_host="DH10B"  
/clone\_lib="Bos taurus ovary fetus"  
/note="Vector: pZLI; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"

## ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 419;  
Best Local Similarity 94.1%; Pred. No. 8.7e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCCCTCTTCGATAC 18  
 |||||  
 DB 192 CCCCCCTCTTCGATGC 208

RESULT 70  
 AZ241055 432 bp DNA linear GSS 15-JUN-2000  
 LOCUS AZ241055/c  
 DEFINITION RPCI-23-72N15.TV RPCI-23 Mus musculus genomic clone RPCI-23-72N15,  
 genomic survey sequence.

ACCESSION AZ241055  
 VERSION AZ241055.1 GI:8554246  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
 1 (bases 1 to 432)  
 Zhao, S., Niernan, W., Feldblum, T., Malek, J., Shatsman, S.,  
 Akinret, B., Levins, M., McGann, S., Teegaye, G., Geer, K., Krol, M., de  
 Jong, P. and Frazer, C.M.  
 Mouse BAC End Sequences from Library RPCI-23  
 Unpublished (1999)  
 Other GSSs: RPCI-23-72N15.TV

TITLE  
 JOURNAL  
 COMMENT  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)  
 or from Resea ch Genetics ([info@resgen.com](mailto:info@resgen.com)). BAC end page:  
[http://www.tigr.org/tdb/bac/ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html)  
 plate: 72 row: N column: 15  
 Seq primer: SP6  
 Class: BAC ends.

FEATURES  
 source  
 Location/Qualifiers

1..432  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-72N15"  
 /sex="Female"  
 /lab\_host="DH10B"  
 /clone\_lib="RPCI-23"  
 /note="Organ: Kidney/Brain; Vector: pBAC3.6; Site\_1:  
 EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBAC3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."

# ORIGIN

Query Match 85.6%; Score 15.4; DB 9; Length 432;  
 Best Local Similarity 94.1%; Pred. No. 8.7e+03;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CCCCCCTCTTCGATA 17  
 |||||  
 DB 239 CCCACCTCTTCGATA 223

Search completed: May 5, 2006, 15:45:31  
 Job time : 2337 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 5, 2006, 15:04:58 ; Search time 96 Seconds  
(without alignments)  
333.293 Million cell updates/sec

Title: US-10-659-980A-5  
Perfect score: 18  
Sequence: 1 cccccccttcgcatcac 18

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues  
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 70 summaries

Database :  
1: /cgn2\_6/ptodata/1/ina/1.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PCRTUS.COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP.COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE.COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	88.9	289	US-10-002-623-449	Sequence 449, App
2	15.4	85.6	6388	US-09-573-080A-179	Sequence 179, App
3	15.4	85.6	20099	US-09-949-016-13784	Sequence 13784, A
4	15.4	85.6	35489	US-09-949-002-731	Sequence 731, App
5	15.4	85.6	35489	US-09-949-002-732	Sequence 731, App
6	15.4	85.6	47915	US-09-949-002-750	Sequence 750, App
7	15.4	85.6	47915	US-09-949-002-751	Sequence 751, App
8	15.4	85.6	51810	US-09-949-002-752	Sequence 752, App
9	15.4	85.6	51810	US-09-949-002-753	Sequence 753, App
10	15	83.3	601	US-09-949-016-30038	Sequence 30038, A
11	15	83.3	601	US-09-949-016-30039	Sequence 30039, A
12	15	83.3	601	US-09-949-016-167012	Sequence 167012, A
13	15	83.3	601	US-09-949-016-167013	Sequence 167013, A
14	15	83.3	2776	US-10-104-047-1147	Sequence 1147, App
15	15	83.3	5581	US-10-164-595-119	Sequence 119, Appl
16	15	83.3	14574	US-09-949-016-16439	Sequence 16439, A
17	15	83.3	46492	US-09-949-016-12953	Sequence 12953, A
18	15	83.3	46492	US-09-949-016-12954	Sequence 12954, A
19	14.8	82.2	601	US-09-949-016-70093	Sequence 70093, A
20	14.8	82.2	601	US-09-949-016-70094	Sequence 70094, A
21	14.8	82.2	26416	US-09-949-016-14266	Sequence 14266, A
22	14.8	82.2	26720	US-09-341-587-7	Sequence 7, Appli
23	14.8	82.2	68283	US-09-949-016-12261	Sequence 12261, A
24	14.8	82.2	97376	US-09-949-016-16093	Sequence 16093, A

ALIGNMENTS

25	14.8	82.2	155617	3	US-09-949-016-16191	Sequence 16191, A
26	14.8	82.2	168971	3	US-09-949-016-13807	Sequence 13807, A
27	14.4	80.0	30	2	US-08-442-809A-74	Sequence 74, Appl
28	14.4	80.0	244	3	US-09-621-976-16605	Sequence 16605, A
29	14.4	80.0	438	3	US-09-513-999C-1993	Sequence 1993, Ap
30	14.4	80.0	601	3	US-09-949-016-25890	Sequence 25890, A
31	14.4	80.0	601	3	US-09-949-016-25891	Sequence 25891, A
32	14.4	80.0	601	3	US-09-949-016-29064	Sequence 29064, A
33	14.4	80.0	601	3	US-09-949-016-61320	Sequence 61320, A
34	14.4	80.0	601	3	US-09-949-016-61321	Sequence 61321, A
35	14.4	80.0	601	3	US-09-949-016-69597	Sequence 69597, A
36	14.4	80.0	601	3	US-09-949-016-171016	Sequence 171016, A
37	14.4	80.0	601	3	US-09-949-016-171017	Sequence 171017, A
38	14.4	80.0	1188	3	US-10-055-001B-12	Sequence 12, Appl
39	14.4	80.0	2028	3	US-09-221-017B-1201	Sequence 1001, Ap
40	14.4	80.0	2325	3	US-09-949-016-562	Sequence 562, App
41	14.4	80.0	2381	2	US-08-021-608D-9	Sequence 9, Appli
42	14.4	80.0	2381	2	US-08-726-160D-9	Sequence 9, Appli
43	14.4	80.0	2384	6	PCT-US94-01782-9	Sequence 1, Appli
44	14.4	80.0	2384	2	US-08-021-608D-1	Sequence 1, Appli
45	14.4	80.0	2384	2	US-08-726-160-1	Sequence 1, Appli
46	14.4	80.0	2384	6	PCT-US94-01782-1	Sequence 2045, Ap
47	14.4	80.0	2475	3	US-09-949-016-2045	Sequence 804, App
48	14.4	80.0	4108	3	US-09-883-096-1	Sequence 13961, A
49	14.4	80.0	6401	3	US-09-221-017B-804	Sequence 13961, A
50	14.4	80.0	13146	3	US-09-949-016-13961	Sequence 12304, A
51	14.4	80.0	34628	3	US-09-949-016-12304	Sequence 13787, A
52	14.4	80.0	34779	3	US-09-949-016-13787	Sequence 16589, A
53	14.4	80.0	36532	3	US-09-949-016-16585	Sequence 12144, A
54	14.4	80.0	47199	3	US-09-949-016-12144	Sequence 13526, A
55	14.4	80.0	47200	3	US-09-949-016-13526	Sequence 17068, A
56	14.4	80.0	89584	3	US-09-949-016-17068	Sequence 13118, A
57	14.4	80.0	107800	3	US-09-949-016-13118	Sequence 11809, A
58	14.4	80.0	116425	3	US-09-949-016-11809	Sequence 11809, A
59	14.4	80.0	300598	3	US-09-949-016-11868	Sequence 14589, A
60	14.4	80.0	302604	3	US-09-949-016-14588	Sequence 14589, A
61	14.4	80.0	302604	3	US-09-949-016-14589	Sequence 17119, A
62	14.4	80.0	308362	3	US-09-949-016-17119	Sequence 1, Appli
63	14.4	80.0	1230025	3	US-09-438-452A-1	Sequence 1, Appli
64	14.4	80.0	1230230	3	US-09-198-452A-1	Sequence 1046, Ap
65	14	77.8	595	3	US-09-513-999C-1046	Sequence 16051, A
66	14	77.8	601	3	US-09-949-016-126051	Sequence 150203, A
67	14	77.8	601	3	US-09-949-016-150203	Sequence 3890, Ap
68	14	77.8	942	3	US-09-513-999C-3890	Sequence 48, Appli
69	14	77.8	1200	3	US-09-222-938A-48	Sequence 930, App
70	14	77.8	4565	3	US-09-949-016-930	

RESULT 1  
US-10-002-623-449/C  
Sequence 449, App Application US/10002623  
Patent No. 692951  
GENERAL INFORMATION: PETER J.  
APPLICANT: UNDERHILL, PETER A.  
TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN  
FILE REFERENCE: STAN-212  
CURRENT APPLICATION NUMBER: US/10/002,623  
PRIOR FILING DATE: 2001-11-01  
PRIOR APPLICATION NUMBER: US 60/245,355  
NUMBER OF SEQ ID NOS: 952  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 449  
LENGTH: 289  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-10-002-623-449

Query Match 88.9%; Score 16; DB 3; Length 289;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCCTCTCTGGATA 17  
|||  
Db 281 CCCCTCTCTGGATA 266

RESULT 2  
US-09-573-080A-179/c  
; Sequence 179, Application US/09573080A  
; Patent No. 6828097  
; GENERAL INFORMATION:  
; APPLICANT: JOAN, KNOLL  
; APPLICANT: ROGAN, PETER  
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI  
; FILE REFERENCE: 30307  
; CURRENT APPLICATION NUMBER: US/09/573,080A  
; CURRENT FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 479  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 179  
; LENGTH: 6388  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: repeat region  
; LOCATION: (1)..(6388)  
; OTHER INFORMATION: mer41  
; NAME/KEY: misc feature  
; OTHER INFORMATION: n is a, c, g or t  
; PUBLICATION INFORMATION:  
; AUTHORS: Jurka, J; Malchiewicz, J; Miosavljivic, A  
; TITLE: Prototypic sequences for human repetitive DNA  
; JOURNAL: Journal of Molecular Evolution  
; VOLUME: 35  
; ISSUE: 4  
; PAGES: 286-291  
; DATE: 1992-10-  
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)  
; DATABASE ENTRY DATE: 1996-01-26  
; US-09-573-080A-179

Query Match 85.6%; Score 15.4; DB 3; Length 6388;  
Best Local Similarity 94.1%; Pred. No. 2.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCTCTCTGGATA 17  
|||  
Db 4010 CCCCTCTCTGGATA 3994

RESULT 3  
US-09-949-016-13784  
; Sequence 13784, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13784  
; LENGTH: 20099  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)....(20099)  
; OTHER INFORMATION: n = A,T,C or G  
; US-09-949-016-13784

Query Match 85.6%; Score 15.4; DB 3; Length 20099;  
Best Local Similarity 94.1%; Pred. No. 3.2e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCCTCTCTGGATAC 18  
|||  
Db 1638 CCCCTCTCTGGATAC 1654

RESULT 4  
US-09-949-002-731  
; Sequence 731, Application US/09949002  
; Patent No. 6900016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION  
; FILE REFERENCE: CL000790  
; CURRENT APPLICATION NUMBER: US/09/949,002  
; CURRENT FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/231,401  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 10823  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 731  
; LENGTH: 35489  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-949-002-731

Query Match 85.6%; Score 15.4; DB 3; Length 35489;  
Best Local Similarity 94.1%; Pred. No. 3.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCTCTCTGGATA 17  
|||  
Db 6486 CCCCTCTCTGGATA 6502

RESULT 5  
US-09-949-002-732  
; Sequence 732, Application US/09949002  
; Patent No. 6900016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION  
; FILE REFERENCE: CL000790  
; CURRENT APPLICATION NUMBER: US/09/949,002  
; CURRENT FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/231,401  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 10823  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 732  
; LENGTH: 35489  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-949-002-732



```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30038
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-30038

```

```

Query Match      83.3%; Score 15; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 CCCCCCTTCTGGA 15
         |||||
Db      256 CCCCCCTTCTGGA 270

```

```

RESULT 11
US-09-949-016-30039
; Sequence 30039, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30039
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-30039

```

```

Query Match      83.3%; Score 15; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 CCCCCCTTCTGGA 15
         |||||
Db      540 CCCCCCTTCTGGA 554

```

```

RESULT 12
US-09-949-016-167012
; Sequence 167012, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167012
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-167012

```

```

Query Match      83.3%; Score 15; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 CCCCCCTTCTGGA 15
         |||||
Db      256 CCCCCCTTCTGGA 270

```

```

RESULT 13
US-09-949-016-167013
; Sequence 167013, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167013
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-167013

```

```

Query Match      83.3%; Score 15; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 CCCCCCTTCTGGA 15
         |||||
Db      540 CCCCCCTTCTGGA 554

```

```

RESULT 14
US-10-104-047-1147/c
; Sequence 1147, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1147
; LENGTH: 2776
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1147

```

```

Query Match      83.3%; Score 15; DB 3; Length 2776;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;

```



Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CCCCCTCTTGATAC 18  
|||||  
Db 2173 CCCCCTCTTGATAC 2159

## RESULT 15

US-10-164-595-19  
; Sequence 19; Application US/10164595  
; Patent No. 6657054  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies, Inc  
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides  
; FILE REFERENCE: 1U 103 R1  
; CURRENT APPLICATION NUMBER: US/10/164,595  
; CURRENT FILING DATE: 2002-06-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 5581  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (264)..(3332)  
; OTHER INFORMATION:  
US-10-164-595-19

Query Match 83.3%; Score 15; DB 3; Length 5581;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCCCTCTTGAT 16  
|||||  
Db 2348 CCCCCTCTTGAT 2362

## RESULT 16

US-09-949-016-16439/C  
; Sequence 16439; Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16439  
; LENGTH: 14574  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)....(14574)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16439

Query Match 83.3%; Score 15; DB 3; Length 14574;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCCTCTTGGA 15  
|||||

Db 2394 CCCCCTCTTGGA 2380

## RESULT 17

US-09-949-016-12953  
; Sequence 12953; Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12953  
; LENGTH: 46492  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)....(46492)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12953

Query Match 83.3%; Score 15; DB 3; Length 46492;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CCCCCTCTTGATA 17  
|||||  
Db 18631 CCCCCTCTTGATA 18645

## RESULT 18

US-09-949-016-12954  
; Sequence 12954; Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12954  
; LENGTH: 46492  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)....(46492)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12954

Query Match 83.3%; Score 15; DB 3; Length 46492;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCCCTCTTGTGATA 17  
| | | | | | | | | |  
| | | | | | | | | |  
Db 18631 CCCCTCTTGTGATA 18645

RESULT 19  
US-09-949-016-70093/c  
; Sequence 70093, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 70093  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-70093

Query Match 82.2%; Score 14.8; DB 3; Length 601;  
Best Local Similarity 88.9%; Pred. No. 3.6e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGTGATAC 18  
| | | | | | | | | |  
| | | | | | | | | |  
Db 600 CCCCCCTCTTGTGATAC 583

RESULT 20  
US-09-949-016-70094/c  
; Sequence 70094, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR FILING DATE: 2000-10-03  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 70094  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-70094

Query Match 82.2%; Score 14.8; DB 3; Length 601;  
Best Local Similarity 88.9%; Pred. No. 3.6e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGTGATAC 18  
| | | | | | | | | |  
| | | | | | | | | |  
Db 101 CCCCCCTCTTGTGATAC 84

RESULT 21  
US-09-949-016-14266/c  
; Sequence 14266, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14266  
; LENGTH: 26416  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14266

Query Match 82.2%; Score 14.8; DB 3; Length 26416;  
Best Local Similarity 88.9%; Pred. No. 6.9e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGTGATAC 18  
| | | | | | | | | |  
| | | | | | | | | |  
Db 24633 CCCCCCTCTTGTGATAC 24616

RESULT 22  
US-09-341-587-7/c  
; Sequence 7, Application US/09341587  
; Patent No. 6346606  
; GENERAL INFORMATION:  
; APPLICANT: Mollenhauer, Jan  
; TITLE OF INVENTION: Protein Containing an SRCR Domain  
; FILE REFERENCE: 4121-108  
; CURRENT APPLICATION NUMBER: US/09/341,587  
; CURRENT FILING DATE: 1999-08-31  
; EARLIER APPLICATION NUMBER: PCT/DE98/00096  
; EARLIER FILING DATE: 1998-01-09  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 28720  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-341-587-7

Query Match 82.2%; Score 14.8; DB 3; Length 28720;  
Best Local Similarity 88.9%; Pred. No. 7e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGTGATAC 18  
| | | | | | | | | |  
| | | | | | | | | |  
Db 1691 CCGCCTCTTGTGATAC 1674

RESULT 23  
US-09-949-016-12261/c  
; Sequence 12261, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12261
; LENGTH: 68283
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12261
```

```
Query Match      82.2% Score 14.8; DB 3; Length 68283;
Best Local Similarity 88.9%; Pred. No. 8.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTCTCTGGATAC 18
Db      26632 CCTCCCTCTCTGGATCC 26615
```

```
RESULT 24
US-09-949-016-16093
; Sequence 16093, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16093
; LENGTH: 97376
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(97376)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-16093
```

```
Query Match      82.2% Score 14.8; DB 3; Length 97376;
Best Local Similarity 88.9%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTCTCTGGATAC 18
Db      565 CACCCCTCTCTGTATAC 582
```

```
RESULT 25
US-09-949-016-16191/C
; Sequence 16191, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16191
; LENGTH: 155617
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16191
```

```
Query Match      82.2% Score 14.8; DB 3; Length 155617;
Best Local Similarity 88.9%; Pred. No. 9.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTCTCTGGATAC 18
Db      4461 CCCCCCTCTCTGAATTC 4444
```

```
RESULT 26
US-09-949-016-13807/C
; Sequence 13807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13807
; LENGTH: 168971
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(168971)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-13807
```

```
Query Match      82.2% Score 14.8; DB 3; Length 168971;
Best Local Similarity 88.9%; Pred. No. 9.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTCTCTGGATAC 18
Db      98800 CCCCCCTTTTGGATAC 98783
```

```
RESULT 27
US-08-442-809A-74/C
; Sequence 74, Application US/08442809A
; Patent No. 5976873
; GENERAL INFORMATION:
; APPLICANT: Bohinski, Robert J.,
; APPLICANT: Whitsett, Jeffrey A.
; TITLE OF INVENTION: Nucleic Acid Sequences
; TITLE OF INVENTION: Controlling Lung Cell -
; TITLE OF INVENTION: Specific Gene Expression
; NUMBER OF SEQUENCES: 76
```

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Carella, Byrne, Bain, Giffillan,  
 ADDRESSEE: Cecchi, Stewart & Olstein  
 STREET: 6 Becker Farm Road  
 CITY: Roseland  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07068  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch diskette  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: WordPerfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/442,809A  
 FILING DATE: 17-MAY-1995  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/245,356  
 FILING DATE: 18-MAY-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Olstein, Elliot M.  
 REGISTRATION NUMBER: 24,025  
 REFERENCE/DOCKET NUMBER: 271010-360  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201-994-1700  
 TELEFAX: 201-994-1744  
 INFORMATION FOR SEQ ID NO: 74:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 30 bases  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: oligonucleotide  
 US-08-442-809A-74

Query Match 80.0%; Score 14.4; DB 2; Length 30;  
 Best Local Similarity 93.8%; Pred. No. 3.4e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCCCTCTTGATGATA 17  
 DB 24 CCCCCCTCTTGATGATA 9  
 RESULT 28  
 US-09-621-976-16605  
 ; Sequence 16605; Application US/09621976  
 ; Patent No. 6639063  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Jobert, S.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: ESTE and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET.054PR2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21  
 ; NUMBER OF SEQ ID NOS: 19335  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 16605  
 ; LENGTH: 244  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-621-976-16605

Query Match 80.0%; Score 14.4; DB 3; Length 244;  
 Best Local Similarity 93.8%; Pred. No. 4.9e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATGAT 16  
 DB 122 CCCCCCTCTTGATGAT 137

RESULT 29  
 US-09-513-999C-1993  
 ; Sequence 1993; Application US/09513999C  
 ; Patent No. 6783961  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Duclet, A.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 ; Patent No. 6783961  
 ; FILE REFERENCE: 59. US2. REG  
 ; CURRENT APPLICATION NUMBER: US/09/513,999C  
 ; CURRENT FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/122,487  
 ; PRIOR FILING DATE: 1999-02-26  
 ; NUMBER OF SEQ ID NOS: 36681  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 1993  
 ; LENGTH: 438  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 107..436  
 US-09-513-999C-1993

Query Match 80.0%; Score 14.4; DB 3; Length 438;  
 Best Local Similarity 93.8%; Pred. No. 5.5e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATGAT 16  
 DB 131 CCCCCCTCTTGATGAT 146  
 RESULT 30  
 US-09-949-016-25890  
 ; Sequence 25890; Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: C1001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 25890  
 ; LENGTH: 601  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-25890

Query Match 80.0%; Score 14.4; DB 3; Length 601;  
 Best Local Similarity 93.8%; Pred. No. 5.8e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTCTTGATGAT 16  
 DB 519 CCCCCCTCTTGATGAT 534  
 RESULT 31  
 US-09-949-016-25891  
 ; Sequence 25891; Application US/09949016

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25891
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-25891
```

```
Query Match          80.0%; Score 14.4; DB 3; Length 601;
Best Local Similarity 93.8%; Pred. No. 5.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CCCCCCTTCTGTGAT 16
Db 160 CCCCCCACTTGTGAT 175
```

```
RESULT 32
US-09-949-016-29064/c
; Sequence 29064, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29064
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-29064
```

```
Query Match          80.0%; Score 14.4; DB 3; Length 601;
Best Local Similarity 93.8%; Pred. No. 5.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CCCCCCTTCTGTGAT 16
Db 326 CCCCCCTTCTGTGAT 311
```

```
RESULT 33
US-09-949-016-61320
; Sequence 61320, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61320
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-61320
```

```
Query Match          80.0%; Score 14.4; DB 3; Length 601;
Best Local Similarity 93.8%; Pred. No. 5.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CCCCCCTTCTGTGAT 16
Db 519 CCCCCCACTTGTGAT 534
```

```
RESULT 34
US-09-949-016-61321
; Sequence 61321, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61321
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-61321
```

```
Query Match          80.0%; Score 14.4; DB 3; Length 601;
Best Local Similarity 93.8%; Pred. No. 5.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CCCCCCTTCTGTGAT 16
Db 160 CCCCCCACTTGTGAT 175
```

```
RESULT 35
US-09-949-016-69597/c
; Sequence 69597, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```

: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 69597
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-69597

```

Query Match	80.0%	Score 14.4	DB 3	Length 601
Best Local Similarity	93.8%	Pred. No. 5	8e+02	
Matches 15	Conservative 0	Mismatches 1	Indels 0	Gaps 0

```

Oy      1 CCCCCCTCTCTGGAT 16
         |||||
Db      326 CCCCCCTCTCTGGCT 311

```

```

: RESULT 36
: US-09-949-016-171016
: Sequence 171016, Application US/09949016
: Patent No. 6812319
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 171016
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-171016

```

Query Match	80.0%	Score 14.4	DB 3	Length 601
Best Local Similarity	93.8%	Pred. No. 5.8e+02		
Matches 15	Conservative	0	Mismatches 1	Indels 0
			Gaps	0

Qy 3 CCCCTCTTCTGGATAC 18  
Db 336 CCCCTCTTCTGGAAC 351

```

RESULT 37
US-09-949-016-171017
; Sequence 171017, Application US/09949016
; Patent No. 682339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

```

```

; SOFTWARE: FASTSQ for Windows Version 4.0
; SEQ ID NO 171017
;
; LENGTH: 601
;
; TYPE: DNA
; ORGANISM: Human
;
US-09-949-016-171017

```

Query Match	80.0%;	Score 14.4;	DB 3;	Length 601;
Best Local Similarity	93.8%;	Pred. No. 5.8e+02;		
Matches 15;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0

```

QY      3 CCCCTCTCTGATAC 18
          |||||
Db      585 CCCCTCTCTGAAC 600

```

```

RESULT 38
US-10-055-001B-12
: Sequence 12, Application US/10055001B
: Patent No. 6933146
: GENERAL INFORMATION:
: APPLICANT: Wesley, Susan V.
: APPLICANT: Waterhouse, Peter
: APPLICANT: Helliwell, Christopher A.
: TITLE OF INVENTION: Methods and means for producing efficient silencing
: TITLE OF INVENTION: constructs using recombinational cloning
: FILE REFERENCE: 021565-108
: CURRENT APPLICATION NUMBER: US/10/055,001B
: CURRENT FILING DATE: 2002-01-25
: PRIOR APPLICATION NUMBER: US 60/264,067
: PRIOR FILING DATE: 2001-01-26
: PRIOR APPLICATION NUMBER: US 60/333,743
: PRIOR FILING DATE: 2001-11-29
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 12
: LENGTH: 1188
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: cDNA sequence of the Arabidopsis thaliana chalcone synthase codi
US-10-055-001B-12

```

Query Match	80.0%;	Score 14.4;	DB 3;	Length 1188;
Best Local Similarity	93.8%;	Pred. No. 6.5e+02;		
Matches 15;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0

Qy 2 CCCCCCTCTCTGGATA 17  
| | | | | | | | | |  
Db 906 CTCCTCTTCTGGATA 921

RESULT 39  
US-09-221-017B-1001  
; Sequence 1001, Application US/09221017B  
; Patent No. 6444799  
; GENERAL INFORMATION:  
; APPLICANT: Rose, Bruce C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: MONROY, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 1001:
SEQUENCE CHARACTERISTICS:
LENGTH: 2028 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc feature
LOCATION: 1...2028
US-09-221-017B-1001

Query Match      80.0%; Score 14.4; DB 3; Length 2028;
Best Local Similarity 93.8%; Pred. No. 7.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 CCCCTCTTGATAC 18
DB      1824 CCCTCTTGATAC 1839

RESULT 40
US-09-949-016-562
Sequence 562, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 562
LENGTH: 2325
TYPE: DNA
ORGANISM: Human
US-09-949-016-562
```

```

Query Match      80.0%; Score 14.4; DB 3; Length 2325;
Best Local Similarity 93.8%; Pred. No. 7.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CCCCTCTTGAT 16
DB      51 CCCCTCTTGACT 66

RESULT 41
US-08-021-608D-9
Sequence 9, Application US/08021608D
Patent No. 5580760
GENERAL INFORMATION:
APPLICANT: LEVENS, DAVID L., DUNCAN,
ROBERT C., AND AVIGAN, MARK I.
TITLE OF INVENTION: NOVEL FUSE BINDING
PROTEIN AND CDNA THEREFOR
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,608D
FILING DATE: 22-FEB-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAM S. FELLER
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2381
TYPE: Nucleic acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: YES
ORIGINAL SOURCE:
ORGANISM: Human
CELL LINE: HL60
FEATURE:
OTHER INFORMATION: 470 bp variable
region where R is A or G.
US-08-021-608D-9

Query Match      80.0%; Score 14.4; DB 2; Length 2381;
Best Local Similarity 93.8%; Pred. No. 7.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CCCCTCTTGAT 16
DB      51 CCCCTCTTGACT 66
```

RESULT 42  
US-08-726-160-9  
; Sequence 9, Application US/08726160  
; Patent No. 5734016  
; GENERAL INFORMATION:  
; APPLICANT: LEVENS, DAVID L., DUNCAN,  
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.  
; TITLE OF INVENTION: NOVEL FUSE BINDING  
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & PINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/726,160  
; FILING DATE: 04-OCT-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/021,608  
; FILING DATE: 22-FEB-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILLIAM S. FEILER  
; REGISTRATION NUMBER: 26,728  
; REFERENCE/DOCKET NUMBER: 2026-40630S1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2381  
; TYPE: Nucleic acid  
; STRANDEDNESS: Double  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: CDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; CELL LINE: HL60  
; FEATURE:  
; OTHER INFORMATION: 470 bp variable  
; OTHER INFORMATION: region where R is A or G.  
US-08-726-160-9  
Query Match 80.0%; Score 14.4; DB 2; Length 2381;  
Best Local Similarity 93.8%; Pred. No. 7.4e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCCCCCTTCTGGAT 16  
Db 51 CCCCCCTTCTGGCT 66

RESULT 43  
PCT-US94-01782-9  
; Sequence 9, Application PC/TUS9401782  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES  
; APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF  
; APPLICANT: HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: NOVEL FUSE BINDING  
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR  
; NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & PINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/01782  
FILING DATE: 22-FEB-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.S. 08/021,608  
FILING DATE: 22-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: WILLIAM S. FEILER  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4063PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2381  
TYPE: Nucleic acid  
STRANDEDNESS: Double  
TOPOLOGY: Unknown  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: YES  
ORIGINAL SOURCE:  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE: HL60  
ORGANELLE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: 470 bp variable  
OTHER INFORMATION: region where R is A or G.  
PCT-US94-01782-9  
Query Match 80.0%; Score 14.4; DB 6; Length 2381;  
Best Local Similarity 93.8%; Pred. No. 7.4e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCCCCCTTCTGGAT 16  
Db 51 CCCCCCTTCTGGCT 66

RESULT 44  
US-08-021-608D-1  
; Sequence 1, Application US/08021608D  
; Patent No. 5580760  
; GENERAL INFORMATION:  
; APPLICANT: LEVENS, DAVID L., DUNCAN,  
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.



```

; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,608D
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-6849
; TELEFAX: (212) 758-4800
; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2384
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; FEATURE:
; OTHER INFORMATION: 473 bp
; OTHER INFORMATION: variable region where R is A or G.
;
; US-08-021-608D-1
;
; Query Match 80.0%; Score 14.4; DB 2; Length 2384;
; Best Local Similarity 93.8%; Pred. No. 7.4e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY 1 CCCCCCTTCTGGAT 16
; Db 51 CCCCCCTTCTGGCT 66
;
; RESULT 45
; US-08-726-160-1
; Sequence 1, Application US/08726160
; Patent No. 5734016
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,160
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/021,608
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-6849
; TELEFAX: (212) 758-4800
; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2384
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; FEATURE:
; OTHER INFORMATION: 473 bp
; OTHER INFORMATION: variable region where R is A or G.
;
; US-08-726-160-1
;
; Query Match 80.0%; Score 14.4; DB 2; Length 2384;
; Best Local Similarity 93.8%; Pred. No. 7.4e+02;
; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
; QY 1 CCCCCCTTCTGGAT 16
; Db 51 CCCCCCTTCTGGCT 66
;
; RESULT 46
; PCT-US94-01782-1
; Sequence 1, Application PC/TUS9401782
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES
; APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
; APPLICANT: HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01782
; FILING DATE: 22-FEB-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```



APPLICATION NUMBER: PP1182  
FILING DATE: 31-DEC-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1546  
FILING DATE: 30-JAN-1998  
PRIOR APPLICATION DATA: PP2911  
FILING DATE: 09-APR-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU98/01023  
FILING DATE: 10-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Montoy, Gladys H  
REGISTRATION NUMBER: 32,430  
REFERENCE/DOCKET NUMBER: 27340-20021.00  
TELEPHONE: 650-813-5600  
TELEFAX: 650-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 804:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6401 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: UNKNOWN  
ORIGINAL SOURCE:  
ORGANISM: PORYPHYROMONAS GINGIVALIS  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1...6401  
US-09-221-017B-804

Query Match 80.0%; Score 14.4; DB 3; Length 6401;  
Best Local Similarity 93.8%; Pred. No. 8.7e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGAT 16  
DB 4764 CCCCCCTTCTTGAT 4749

RESULT 50  
US-09-949-016-13961  
Sequence 13961, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13961  
LENGTH: 13146  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-13961

Query Match 80.0%; Score 14.4; DB 3; Length 13146;  
Best Local Similarity 93.8%; Pred. No. 9.9e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGAT 16  
DB 5403 CCCCCCTTCTTGAT 5418

RESULT 51  
US-09-949-016-12304  
Sequence 12304, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12304  
LENGTH: 34628  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-12304

Query Match 80.0%; Score 14.4; DB 3; Length 34628;  
Best Local Similarity 93.8%; Pred. No. 1.2e+03;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGAT 16  
DB 2051 CCCCCCTTCTTGAT 2066

RESULT 52  
US-09-949-016-13787  
Sequence 13787, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13787  
LENGTH: 34779  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-13787

Query Match 80.0%; Score 14.4; DB 3; Length 34779;  
Best Local Similarity 93.8%; Pred. No. 1.2e+03;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGAT 16  
DB 2051 CCCCCCTTCTTGAT 2066

## RESULT 53

```
US-09-949-016-16585/c
; Sequence 16585, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16585
; LENGTH: 36532
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(36532)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16585
```

```
Query Match          80.0%; Score 14.4; DB 3; Length 36532;
Best Local Similarity 93.8%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

## QY 3 CCCCTCTTGTGAT 18

```
Db 23533 CCCCTCTTGTGAAAC 23518
```

## RESULT 54

```
US-09-949-016-12144
; Sequence 12144, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12144
; LENGTH: 47199
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12144
```

```
Query Match          80.0%; Score 14.4; DB 3; Length 47199;
Best Local Similarity 93.8%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

## QY 1 CCCCTCTTGTGAT 16

```
Db 23860 CCCCTCTTGTGAT 23875
```

## RESULT 55

## US-09-949-016-13526

```
; Sequence 13526, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13526
; LENGTH: 47200
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13526
```

```
Query Match          80.0%; Score 14.4; DB 3; Length 47200;
Best Local Similarity 93.8%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

## QY 1 CCCCTCTTGTGAT 16

```
Db 23860 CCCCTCTTGTGAT 23875
```

## RESULT 56

```
US-09-949-016-17068/c
; Sequence 17068, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17068
; LENGTH: 89584
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17068
```

```
Query Match          80.0%; Score 14.4; DB 3; Length 89584;
Best Local Similarity 93.8%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

## QY 2 CCCCTCTTGTGAT 17

```
Db 28576 CCCCTCTTGTGATA 28561
```

## RESULT 57

```
US-09-949-016-13118
; Sequence 13118, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13118
; LENGTH: 107800
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13118

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 3; Length 107800;
Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCTCTCTTGAT 16
Db 51520 CCCCTCTCTGTAT 51535

RESULT 58
US-09-949-016-11809
; Sequence 11809, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11809
; LENGTH: 116425
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11809

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 3; Length 116425;
Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCTCTCTTGAT 16
Db 60146 CCCCTCTCTGTAT 60161

RESULT 59
US-09-949-016-11868
; Sequence 11868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11868
; LENGTH: 300598
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(300598)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11868

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 3; Length 300598;
Best Local Similarity 93.8%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCCTCTCTGATAC 18
Db 150986 CCCCTCTCTGCTAC 151001

RESULT 60
US-09-949-016-14588
; Sequence 14588, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14588
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14588

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 3; Length 302604;
Best Local Similarity 93.8%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCCTCTCTGATAC 18
Db 200992 CCCCTCTCTGCTAC 201007

RESULT 61
US-09-949-016-14589
; Sequence 14589, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 14589
LENGTH: 302604
TYPE: DNA
ORGANISM: Human
NAME/KEY: misc_feature
LOCATION: (1)...(302604)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14589

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 3; Length 302604;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 CCCCTCTCTGATAC 18
Db 200992 CCCCTCTCTGCTAC 201007

RESULT 62
US-09-949-016-17119
Sequence 17119, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 17119
LENGTH: 308362
TYPE: DNA
ORGANISM: Human
NAME/KEY: misc_feature
LOCATION: (1)...(308362)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17119

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 3; Length 308362;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 CCCCTCTCTGATAC 18
Db 200808 CCCCTCTCTGCTAC 200823

RESULT 63
US-09-198-452A-1
Sequence 1, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
```

```
APPLICANT: Grifflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
thereof and uses thereof, in particular for the diagnosis, prevention
and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO: 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (15001)...(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (30001)...(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (45001)...(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (60001)...(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (75001)...(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (90001)...(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (105001)...(120000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (120001)...(135000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (135001)...(150000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (150001)...(165000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (165001)...(180000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (180001)...(195000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (195001)...(210000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (210001)...(225000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (225001)...(240000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (240001)...(255000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (255001)...(270000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (270001)...(285000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (285001)...(300000)
OTHER INFORMATION: n=a or c or g or t
```

```
NAME/KEY: misc_feature
LOCATION: (300001)..(315000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (315001)..(330000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
```

```
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
```

```
Query Match 80.0%; Score 14.4; DB 3; Length 1230025;
Best local similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0;
```

```
Qy 2 CCCCTCTTGATA 17
Db 936810 CCCGCTCTTGATA 936825
```

```
RESULT 64
US-09-438-185A-1
; Sequence 1, Application US/09438185A
; Patent No. 682071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kaiman, Sue
; APPLICANT: Davis, Ronald
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
```

```

; CURRENT APPLICATION NUMBER: US/09/438,185A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1230230
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-438-185A-1

```

```

Query Match      80.0%; Score 14.4; DB 3; Length 1230230;
Best Local Similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 CCCCCCTCTGTGATA 17
Db      926686 CCCCCCTCTGTGATA 926701

```

```

RESULT 65
US-09-513-999C-1046
; Sequence 1046, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 1046
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 315..593
US-09-513-999C-1046

```

```

Query Match      77.8%; Score 14; DB 3; Length 595;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 CCCCCCTCTGTGGA 15
Db      59 CCCCCCTCTGTGGA 72

```

```

RESULT 66
US-09-949-016-126051/c
; Sequence 126051, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

```

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126051
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-126051

```

```

Query Match      77.8%; Score 14; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 CCCCCCTCTGTGGA 15
Db      529 CCCCCCTCTGTGGA 516

```

```

RESULT 67
US-09-949-016-150203/c
; Sequence 150203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150203
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-150203

```

```

Query Match      77.8%; Score 14; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 CCCCCCTCTGTGAT 16
Db      481 CCCCCCTCTGTGAT 468

```

```

RESULT 68
US-09-513-999C-3890
; Sequence 3890, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 3890
; LENGTH: 942
; TYPE: DNA

```



ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 315..839  
US-09-513-999C-3890

Query Match 77.8%; Score 14; DB 3; Length 942;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCCCTCTCTGGA 15  
|||||  
Db 59 CCCCCTCTCTGGA 72

RESULT 69  
US-09-222-938A-48  
Sequence 48, Application US/09222938A

Patent No. 6437108  
GENERAL INFORMATION:  
APPLICANT: Youngman, Philip  
APPLICANT: Filtz, Christian  
APPLICANT: Murphy, Christopher  
APPLICANT: Guzman, Luz-Maria  
TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE  
FILE REFERENCE: 0734/060001  
CURRENT APPLICATION NUMBER: US/09/222,938A  
CURRENT FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 102  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 48  
LENGTH: 1200  
TYPE: DNA  
ORGANISM: Streptococcus pneumoniae  
US-09-222-938A-48

Query Match 77.8%; Score 14; DB 3; Length 1200;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CCTCTCTGGATAC 18  
|||||  
Db 1180 CCTCTCTGGATAC 1193

RESULT 70  
US-09-949-016-930  
Sequence 930, Application US/09949016

Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C1001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 930  
LENGTH: 4565  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-930

Query Match 77.8%; Score 14; DB 3; Length 4565;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CCTCTCTGGATAC 18  
|||||  
Db 3627 CCTCTCTGGATAC 3640

Search completed: May 5, 2006, 15:47:16  
Job time : 104 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 5, 2006, 16:59:51 ; Search time 497 Seconds  
(without alignments)  
299.495 Million cell updates/sec

Title: US-10-659-980A-5

Perfect score: 18

Sequence: 1 cccccctctcgcgatac 18

Scoring table: IDENTITY NUC  
Gapco 10.0, Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 70 summaries

Database: Published Applications\_NA\_Main:\*

1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*  
5: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
6: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
7: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
8: /cgn2\_6/prodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
9: /cgn2\_6/prodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
10: /cgn2\_6/prodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	18	US-10-659-948A-5	Sequence 5, Appl1
2	18	100.0	18	US-10-659-980A-5	Sequence 5, Appl1
3	18	100.0	18	US-10-659-983A-5	Sequence 5, Appl1
4	18	100.0	1457	US-10-659-948A-1	Sequence 1, Appl1
5	18	100.0	1457	US-10-659-948A-2	Sequence 1, Appl1
6	18	100.0	1457	US-10-659-980A-1	Sequence 1, Appl1
7	18	100.0	1457	US-10-659-980A-2	Sequence 1, Appl1
8	18	100.0	1457	US-10-659-983A-1	Sequence 1, Appl1
9	18	100.0	1457	US-10-659-983A-2	Sequence 1, Appl1
10	16.4	91.1	361	US-10-424-599-49454	Sequence 49454, A
11	16.4	91.1	436	US-10-424-599-126572	Sequence 126572, A
12	16	88.9	289	US-10-002-623-449	Sequence 449, App
13	16	88.9	460	US-10-425-115-133449	Sequence 133449, A
14	15.4	85.6	630	US-09-925-065A-358412	Sequence 358412, A
15	15.4	85.6	1104	US-10-282-122A-15346	Sequence 15346, A
16	15.4	85.6	1899	US-10-424-599-40637	Sequence 40637, A
17	15.4	85.6	3263	US-09-925-065A-723345	Sequence 723345, A
18	15.4	85.6	6388	US-09-854-867-179	Sequence 179, App
19	15.4	85.6	6388	US-10-786-970A-179	Sequence 179, App
20	15.4	85.6	12542	US-11-097-143-15784	Sequence 15784, A
21	15.4	85.6	67810	US-10-087-192-1738	Sequence 1738, App
22	15.4	85.6	79256	US-10-322-281-167	Sequence 167, App
23	15	83.3	240	US-09-783-590-2705	Sequence 2705, App

## ALIGNMENTS

24	15	83.3	333	8	US-10-425-115-25159	Sequence 25159, A
25	15	83.3	409	8	US-09-918-995-16197	Sequence 16197, A
26	15	83.3	442	8	US-10-425-115-180477	Sequence 180477, A
27	15	83.3	499	4	US-09-925-065A-186619	Sequence 186619, A
28	15	83.3	1655	8	US-10-425-115-97068	Sequence 97068, A
29	15	83.3	2776	6	US-10-104-047-1147	Sequence 1147, App
30	15	83.3	5581	9	US-10-717-665-197	Sequence 197, Appl
31	15	83.3	60989	6	US-10-292-798-1197	Sequence 122, App
32	15	83.3	421609	7	US-10-367-094-1122	Sequence 16233, A
33	14.8	82.2	176	7	US-10-242-535A-18239	Sequence 16233, A
34	14.8	82.2	176	7	US-10-085-783A-16239	Sequence 16233, A
35	14.8	82.2	183	7	US-10-242-535A-22879	Sequence 22879, A
36	14.8	82.2	183	7	US-10-085-783A-22879	Sequence 22879, A
37	14.8	82.2	209	7	US-10-242-535A-25752	Sequence 25752, A
38	14.8	82.2	209	7	US-10-085-783A-25752	Sequence 25752, A
39	14.8	82.2	342	8	US-10-425-115-58042	Sequence 58042, A
40	14.8	82.2	413	6	US-10-252-157-147	Sequence 147, App
41	14.8	82.2	433	6	US-10-674-124A-6727	Sequence 6727, App
42	14.8	82.2	437	4	US-10-424-599-116257	Sequence 116257, A
43	14.8	82.2	547	4	US-09-925-065A-651102	Sequence 651102, A
44	14.8	82.2	557	7	US-10-437-963-48133	Sequence 48133, A
45	14.8	82.2	564	7	US-10-437-963-48132	Sequence 48132, A
46	14.8	82.2	566	4	US-09-925-065A-330692	Sequence 330692, A
47	14.8	82.2	591	4	US-09-925-065A-736648	Sequence 736648, A
48	14.8	82.2	591	4	US-09-925-065A-736649	Sequence 736649, A
49	14.8	82.2	591	4	US-09-925-065A-814201	Sequence 814201, A
50	14.8	82.2	591	4	US-09-925-065A-814202	Sequence 814202, A
51	14.8	82.2	604	8	US-10-425-115-139469	Sequence 139469, A
52	14.8	82.2	632	4	US-09-925-065A-953406	Sequence 953406, A
53	14.8	82.2	638	4	US-09-925-065A-953455	Sequence 953455, A
54	14.8	82.2	679	5	US-10-027-632-21681	Sequence 21681, A
55	14.8	82.2	1095	6	US-10-027-632-21681	Sequence 21681, A
56	14.8	82.2	1157	9	US-10-450-763-28191	Sequence 36419, A
57	14.8	82.2	1209	5	US-10-027-632-123726	Sequence 123726, A
58	14.8	82.2	1209	6	US-10-027-632-123726	Sequence 123726, A
59	14.8	82.2	1371	4	US-09-925-065A-53648	Sequence 53648, A
60	14.8	82.2	1371	4	US-09-925-065A-53649	Sequence 53649, A
61	14.8	82.2	1371	4	US-09-925-065A-61730	Sequence 61730, A
62	14.8	82.2	1890	4	US-09-925-065A-61730	Sequence 61730, A
63	14.8	82.2	1890	5	US-10-027-632-261924	Sequence 261924, A
64	14.8	82.2	1890	5	US-10-027-632-261925	Sequence 261925, A
65	14.8	82.2	1890	6	US-10-027-632-261924	Sequence 261924, A
66	14.8	82.2	2253	5	US-10-027-632-110327	Sequence 261925, A
67	14.8	82.2	2253	5	US-10-027-632-110327	Sequence 110327, A
68	14.8	82.2	2297	8	US-10-739-930-5195	Sequence 5195, App
69	14.8	82.2	2433	9	US-10-450-763-14809	Sequence 14809, A
70	14.8	82.2	2433	9	US-10-450-763-14809	Sequence 14809, A

## RESULT 1

US-10-659-948A-5  
; Sequence 5, Application US/10659948A  
; Publication No. US20040101946A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovance, Timothy A  
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-284309  
; CURRENT APPLICATION NUMBER: US/10/659, 948A  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573, 664  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386, 217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386, 218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386, 219  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5

```
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Oligonucleotide Probe
US-10-659-948A-5
```

```
Query Match          100.0%; Score 18; DB 7; Length 18;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTTCGATAC 18
        |||
Db       1 CCCCCCTCTTCGATAC 18
```

## RESULT 2

```
US-10-659-980A-5
; Sequence 5, Application US/10659980A
; Publication No. US2004010613A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneec, Timothy A
; TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284781
; CURRENT APPLICATION NUMBER: US/10/659,980A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Oligonucleotide Probe
US-10-659-980A-5
```

```
Query Match          100.0%; Score 18; DB 7; Length 18;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTTCGATAC 18
        |||
Db       1 CCCCCCTCTTCGATAC 18
```

## RESULT 3

```
US-10-659-983A-5
; Sequence 5, Application US/10659983A
; Publication No. US2004015731A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneec, Timothy A
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284779
; CURRENT APPLICATION NUMBER: US/10/659,983A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 5
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Oligonucleotide Probe
US-10-659-983A-5
```

```
Query Match          100.0%; Score 18; DB 7; Length 18;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTTCGATAC 18
        |||
Db       1 CCCCCCTCTTCGATAC 18
```

## RESULT 4

```
US-10-659-948A-1/c
; Sequence 1, Application US/10659948A
; Publication No. US20040101946A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneec, Timothy A
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-294309
; CURRENT APPLICATION NUMBER: US/10/659,948A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type A R1clone140 16S rDNA
US-10-659-948A-1
```

```
Query Match          100.0%; Score 18; DB 7; Length 1457;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTTCGATAC 18
        |||
Db       118 CCCCCCTCTTCGATAC 101
```

## RESULT 5

```
US-10-659-948A-2/c
; Sequence 2, Application US/10659948A
; Publication No. US20040101946A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneec, Timothy A
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-294309
; CURRENT APPLICATION NUMBER: US/10/659,948A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
```

```
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 2
/ LENGTH: 1457
/ TYPE: DNA
/ ORGANISM: Unknown
/ FEATURE:
/ OTHER INFORMATION: AOB Type A1 R7clone187 16S rDNA
US-10-659-948A-2
```

```
Query Match          100.0%; Score 18; DB 7; Length 1457;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CCCCCCTCTTGTGATAC 18
    |||||
Db 118 CCCCCCTCTTGTGATAC 101
```

```
RESULT 6
US-10-659-980A-1/c
/ Sequence 1, Application US/10659980A
/ Publication No. US20040106133A1
/ GENERAL INFORMATION:
/ APPLICANT: Hovaneec, Timothy A
/ TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria
/ FILE REFERENCE: 81289-284781
/ CURRENT APPLICATION NUMBER: US/10/659,980A
/ PRIOR FILING DATE: 2003-09-10
/ PRIOR APPLICATION NUMBER: US 09/573,684
/ PRIOR FILING DATE: 2000-05-19
/ PRIOR APPLICATION NUMBER: US 60/386,217
/ PRIOR FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: US 60/386,218
/ PRIOR FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: US 60/386,219
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1
/ LENGTH: 1457
/ TYPE: DNA
/ ORGANISM: Unknown
/ FEATURE:
/ OTHER INFORMATION: AOB Type A R7clone140 16S rDNA
US-10-659-980A-1
```

```
Query Match          100.0%; Score 18; DB 7; Length 1457;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CCCCCCTCTTGTGATAC 18
    |||||
Db 118 CCCCCCTCTTGTGATAC 101
```

```
RESULT 7
US-10-659-980A-2/c
/ Sequence 2, Application US/10659980A
/ Publication No. US20040106133A1
/ GENERAL INFORMATION:
/ APPLICANT: Hovaneec, Timothy A
/ TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria
/ FILE REFERENCE: 81289-284781
/ CURRENT APPLICATION NUMBER: US/10/659,980A
/ PRIOR FILING DATE: 2003-09-10
/ PRIOR APPLICATION NUMBER: US 09/573,684
/ PRIOR FILING DATE: 2000-05-19
/ PRIOR APPLICATION NUMBER: US 60/386,217
/ PRIOR FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: US 60/386,218
/ PRIOR FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: US 60/386,219
/ PRIOR FILING DATE: 2002-09-19
```

```
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 2
/ LENGTH: 1457
/ TYPE: DNA
/ ORGANISM: Unknown
/ FEATURE:
/ OTHER INFORMATION: AOB Type A1 R7clone187 16S rDNA
US-10-659-980A-2
```

```
Query Match          100.0%; Score 18; DB 7; Length 1457;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CCCCCCTCTTGTGATAC 18
    |||||
Db 118 CCCCCCTCTTGTGATAC 101
```

```
RESULT 8
US-10-659-983A-1/c
/ Sequence 1, Application US/10659983A
/ Publication No. US20040157313A1
/ GENERAL INFORMATION:
/ APPLICANT: Hovaneec, Timothy A
/ TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
/ FILE REFERENCE: 81289-284779
/ CURRENT APPLICATION NUMBER: US/10/659,983A
/ PRIOR FILING DATE: 2003-09-10
/ PRIOR APPLICATION NUMBER: US 09/573,684
/ PRIOR FILING DATE: 2000-05-19
/ PRIOR APPLICATION NUMBER: US 60/386,217
/ PRIOR FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: US 60/386,218
/ PRIOR FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: US 60/386,219
/ PRIOR FILING DATE: 2002-09-19
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1
/ LENGTH: 1457
/ TYPE: DNA
/ ORGANISM: Unknown
/ FEATURE:
/ OTHER INFORMATION: AOB Type A R7clone140 16S rDNA
US-10-659-983A-1
```

```
Query Match          100.0%; Score 18; DB 7; Length 1457;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CCCCCCTCTTGTGATAC 18
    |||||
Db 118 CCCCCCTCTTGTGATAC 101
```

```
RESULT 9
US-10-659-983A-2/c
/ Sequence 2, Application US/10659983A
/ Publication No. US20040157313A1
/ GENERAL INFORMATION:
/ APPLICANT: Hovaneec, Timothy A
/ TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
/ FILE REFERENCE: 81289-284779
/ CURRENT APPLICATION NUMBER: US/10/659,983A
/ PRIOR FILING DATE: 2003-09-10
/ PRIOR APPLICATION NUMBER: US 09/573,684
/ PRIOR FILING DATE: 2000-05-19
/ PRIOR APPLICATION NUMBER: US 60/386,217
/ PRIOR FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: US 60/386,218
/ PRIOR FILING DATE: 2002-09-19
/ PRIOR APPLICATION NUMBER: US 60/386,219
```

```

; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type A1 R7clone187 16S rDNA
US-10-659-983A-2

Query Match
Best Local Similarity 100.0%; Score 18; DB 7; Length 1457;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18
Db 118 CCCCCCTTCTGGATAC 101

RESULT 10
US-10-424-599-49454/C
; Sequence 49454, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 49454
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_15664C.1
US-10-424-599-49454

Query Match
Best Local Similarity 91.1%; Score 16.4; DB 7; Length 361;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18
Db 288 CCCCCCTTCTGGATAC 271

RESULT 11
US-10-424-599-126572
; Sequence 126572, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 126572
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
```

```

; LOCATION: (1)..(436)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_85300C.1
US-10-424-599-126572

Query Match
Best Local Similarity 91.1%; Score 16.4; DB 7; Length 436;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGGATAC 18
Db 223 CCCCCCTTCTGGACAC 240

RESULT 12
US-10-002-623-449/C
; Sequence 449, Application US/10002623
; Publication No. US20030134285A1
; GENERAL INFORMATION:
; APPLICANT: OSFNER, PETER J.
; APPLICANT: UNDERHILL, PETER A.
; TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
; FILE REFERENCE: POPULATIONS
; CURRENT APPLICATION NUMBER: US/10/002,623
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,355
; NUMBER OF SEQ ID NOS: 952
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-002-623-449

Query Match
Best Local Similarity 88.9%; Score 16; DB 6; Length 289;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCCCCTTCTGGATA 17
Db 281 CCCCCCTTCTGGATA 266

RESULT 13
US-10-425-115-133349/C
; Sequence 133349, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 133349
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_53102C.1
US-10-425-115-133349

Query Match
Best Local Similarity 100.0%; Score 16; DB 8; Length 460;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 CCCCCCTTCTTGAT 16  
| | | | | | | | | | | | | | | | | | | | | |  
Db 191 CCCCCCTTCTTGAT 176

RESULT 14  
US-09-925-065A-358412/c  
; Sequence 358412, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 358412  
; LENGTH: 630  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-358412

Query Match 85.6%; Score 15.4; DB 4; Length 630;  
Best Local Similarity 94.1%; Pred. No. 6.7e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGATA 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 290 CCCCCCTTCTTGATA 274

RESULT 15  
US-10-282-122A-15346/c  
; Sequence 15346, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: EUTRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15346  
; LENGTH: 1104  
; TYPE: DNA  
; ORGANISM: Clostridium acetobutylicum  
US-10-282-122A-15346

Query Match 85.6%; Score 15.4; DB 7; Length 1104;  
Best Local Similarity 94.1%; Pred. No. 6.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCCCTTCTTGATAC 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 865 CCCCCCTTCTTGATAC 849

RESULT 16  
US-10-424-599-40697/c  
; Sequence 40697, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 40697  
; LENGTH: 1899  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_136749C.1  
US-10-424-599-40697

Query Match 85.6%; Score 15.4; DB 7; Length 1899;  
Best Local Similarity 94.1%; Pred. No. 6.4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGATA 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1287 CCCCCCTTCTTGATA 1271

RESULT 17  
US-09-925-065A-723345  
; Sequence 723345, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08

```

; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 723345
; LENGTH: 3263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-723345
```

```

Query Match      85.6%; Score 15.4; DB 4; Length 3263;
Best Local Similarity 94.1%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1  CCCCCCTTCTCGATA 17
Db      268 CCTCCCTCTTCTGATA 284
```

```

RESULT 18
US-09-854-867-179/c
; Sequence 179, Application US/09854867
; Publication No. US20030224356A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/854,867
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 179
; LENGTH: 6388
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(6388)
; OTHER INFORMATION: mer4i
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (731)..(731)
; OTHER INFORMATION: n is a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (922)..(922)
; OTHER INFORMATION: n is a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2530)..(2530)
; OTHER INFORMATION: n is a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2879)..(2879)
; OTHER INFORMATION: n is a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4632)..(4632)
; OTHER INFORMATION: n is a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6128)..(6128)
; OTHER INFORMATION: n is a, c, g or t
US-09-854-867-179
```

```

Query Match      85.6%; Score 15.4; DB 3; Length 6388;
Best Local Similarity 94.1%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1  CCCCCCTTCTCGATA 17
Db      4010 CCCACTCTTCTGATA 3994
```

```

RESULT 19
US-10-786-970A-179/c
; Sequence 179, Application US/10786970A
; Publication No. US20050064449A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/10/786,970A
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/573,080
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 179
; LENGTH: 6388
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(6388)
; OTHER INFORMATION: mer4i
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n is a, c, g or t
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
; DATABASE ENTRY DATE: 1996-01-26
US-10-786-970A-179
```

```

Query Match      85.6%; Score 15.4; DB 9; Length 6388;
Best Local Similarity 94.1%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

Qy      1  CCCCCCTTCTCGATA 17
Db      4010 CCCACTCTTCTGATA 3994

RESULT 20
US-11-097-143-15784
; Sequence 15784, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
US-11-097-143-15784
```



```
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15784
; LENGTH: 12542
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-15784
```

```
Query Match      85.6%; Score 15.4; DB 10; Length 12542;
Best Local Similarity 94.1%; Pred. No. 6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 CCCCCCTCTCTGGATA 17
          |||||
Db      5329 CCCCCCTCTCTGGATA 5345
```

```
RESULT 21
US-10-087-192-1738
; Sequence 1738, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1738
; LENGTH: 67810
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1738

Query Match      85.6%; Score 15.4; DB 5; Length 67810;
Best Local Similarity 94.1%; Pred. No. 5.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 CCCCCCTCTCTGGATA 17
          |||||
Db      23353 CCCCCTCTCTGGATA 23369
```

```
RESULT 22
US-10-322-281-167/c
; Sequence 167, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
```

```
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167
; LENGTH: 79256
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(79256)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-167
```

```
Query Match      85.6%; Score 15.4; DB 7; Length 79256;
Best Local Similarity 94.1%; Pred. No. 5.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 CCCCCCTCTCTGGATA 17
          |||||
Db      35784 CCCCCTCTCTGGATA 35768
```

```
RESULT 23
US-09-783-590-2705/c
; Sequence 2705, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16,2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2705
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (13)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (27)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (56)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (86)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (149)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (156)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (172)
; OTHER INFORMATION: n equals a,t,g, or c
```

```

; NAME/KEY: misc feature
; LOCATION: (175)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (185)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (186)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (239)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-2705

Query Match      83.3%; Score 15; DB 3; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  CCCCCTCTCTGGAT 16
DB      101 CCCCCTCTCTGGAT 87

RESULT 24
US-10-425-115-25159/c
; Sequence 25159, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 25159
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_12294C.1
US-10-425-115-25159

Query Match      83.3%; Score 15; DB 8; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  CCCCCTCTCTGGAT 16
DB      240 CCCCCTCTCTGGAT 226

RESULT 25
US-09-918-995-16197/c
; Sequence 16197, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hybreg, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16197
; LENGTH: 409
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)....(409)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-16197

Query Match      83.3%; Score 15; DB 3; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  CCCCCTCTCTGGAT 16
DB      168 CCCCCTCTCTGGAT 154

RESULT 26
US-10-425-115-180477/c
; Sequence 180477, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 180477
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_96175C.1
US-10-425-115-180477

Query Match      83.3%; Score 15; DB 8; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CCCCCTCTCTGGA 15
DB      440 CCCCCTCTCTGGA 426

RESULT 27
US-09-925-065A-166619/c
; Sequence 166619, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 166619
; LENGTH: 499
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-166619

Query Match      83.3%; Score 15; DB 4; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCCCCTTCTTGAT 16
      |||
Db      182 CCCCCTTCTTGAT 168

RESULT 28
US-10-425-115-97068
; Sequence 97068, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 97068
; LENGTH: 1655
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_2003C.1
US-10-425-115-97068

Query Match      83.3%; Score 15; DB 8; Length 1655;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CCCCCTTCTTGGA 15
      |||
Db      1472 CCCCCTTCTTGGA 1486

RESULT 29
US-10-104-047-1147/C
; Sequence 1147, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1147
; LENGTH: 2776
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1147

Query Match      83.3%; Score 15; DB 6; Length 2776;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 CCGCTTCTGATAC 18
      |||
Db      2173 CCGCTTCTGATAC 2159
```

```
RESULT 30
US-10-717-665-19
; Sequence 19, Application US/10717665
; Publication No. US20050106579A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: IU 103 R1
; CURRENT APPLICATION NUMBER: US/10/717,665
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/10/164,595
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Version 3.1
; SEQ ID NO 19
; LENGTH: 5581
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (264)..(3332)
; OTHER INFORMATION:
US-10-717-665-19

Query Match      83.3%; Score 15; DB 9; Length 5581;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCCCCTTCTTGAT 16
      |||
Db      2348 CCCCCTTCTTGAT 2362

RESULT 31
US-10-292-798-197/C
; Sequence 197, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: ARIYAMA, YUTAKA
; APPLICANT: ABRARANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 197
; LENGTH: 60989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(60989)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(611)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39988)..(40012)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (60293)..(60789)
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (28886)..(28985)
```

```

; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (29108)..(29108)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (31447)..(31546)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (34276)..(34375)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (36877)..(36976)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (40482)..(40581)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (42087)..(42087)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (44965)..(45064)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (45633)..(45633)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (49080)..(49179)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (52751)..(52850)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (52920)..(52920)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (57435)..(57534)
; OTHER INFORMATION: a, t, c, g, unknown or other
US-10-292-798-197

Query Match      83.3%; Score 15; DB 6; Length 60989;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4  CCCTCTTGTGGATAC 18
          |||||
Db      51268  CCCTCTTGTGGATAC 51254

RESULT 32
US-10-367-094-122
; Sequence 122, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 122
; LENGTH: 421609
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(421609)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-122

Query Match      83.3%; Score 15; DB 7; Length 421609;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  CCCCTCTTGTGGAT 16
          |||||
Db      55136  CCCCTCTTGTGGAT 55150

RESULT 33
US-10-242-535A-16239/C
; Sequence 16239, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16239
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-16239

Query Match      82.2%; Score 14.8; DB 7; Length 176;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CCCCTCTTGTGGATAC 18
          |||||
Db      90  CCCCTCTTGTGGATAC 73

RESULT 34
US-10-085-783A-16239/C
; Sequence 16239, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
```

```
SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16239
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-16239
```

```
Query Match      82.2%; Score 14.8; DB 7; Length 176;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTCTGATAC 18
          ||||| ||||| |||||
Db      90 CCCCCCGCTTCTGGCTAC 73
```

```
RESULT 35
US-10-242-535A-22879/c
; Sequence 22879, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22879
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (17)..(17)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (153)..(153)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-22879
```

```
Query Match      82.2%; Score 14.8; DB 7; Length 183;
```

```
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
```

```
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTCTGATAC 18
          ||||| ||||| |||||
Db      98 CCCCCCGCTTCTGGCTAC 81
```

```
RESULT 36
US-10-085-783A-22879/c
; Sequence 22879, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
```

```
CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22879
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (17)..(17)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (153)..(153)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-22879
```

```
Query Match      82.2%; Score 14.8; DB 7; Length 183;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTCTGATAC 18
          ||||| ||||| |||||
Db      98 CCCCCCGCTTCTGGCTAC 81
```

```
RESULT 37
US-10-242-535A-25752/c
; Sequence 25752, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25752
; LENGTH: 209
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (27)..(27)
```

```
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (42)..(42)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44)..(44)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(51)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (58)..(58)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (78)..(78)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-25752

Query Match      82.2%; Score 14.8; DB 7; Length 209;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CCCCCCTCTTGATAC 18
DB      139 CCCCCGCTTGTGCTAC 122

RESULT 38
US-10-085-783A-25752/c
; Sequence 25752, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,555
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25752
; LENGTH: 209
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (27)..(27)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (42)..(42)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (42)..(42)
; OTHER INFORMATION: n is a, c, g, or t
; NAME/KEY: misc_feature
```

```
; LOCATION: (44)..(44)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(51)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (58)..(58)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (78)..(78)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-25752

Query Match      82.2%; Score 14.8; DB 7; Length 209;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CCCCCCTCTTGATAC 18
DB      139 CCCCCGCTTGTGCTAC 122

RESULT 39
US-10-425-115-58042/c
; Sequence 58042, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 58042
; LENGTH: 342
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(342)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MHT4577_152930C.1
US-10-425-115-58042

Query Match      82.2%; Score 14.8; DB 8; Length 342;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CCCCCCTCTTGATAC 18
DB      184 CCCCCCTCTTGATAC 167

RESULT 40
US-10-252-157-147/c
; Sequence 147, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
```

```
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO: 147
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 404197.8c
US-10-252-157-147
```

```
Query Match      82.2%; Score 14.8; DB 6; Length 413;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTCTTCTGGATAC 18
Db      95 CCCCCCTCTGCTGGCTAC 78
```

```
RESULT 41
US-10-674-124A-6727/c
; Sequence 6727, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidecoshi
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 6727
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: chr4.fa.07fz.31705438
; FEATURE:
; OTHER INFORMATION: Located on chromosome 4
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; FEATURE:
; OTHER INFORMATION: sequence : 31496875
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 92289
US-10-674-124A-6727
```

```
Query Match      82.2%; Score 14.8; DB 8; Length 433;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTCTTCTGGATAC 18
Db      194 CCCCCCACTTCTGATTC 177
```

```
RESULT 42
US-10-424-599-116257
```

```
; Sequence 116257, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 116257
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_75991C.1
US-10-424-599-116257
```

```
Query Match      82.2%; Score 14.8; DB 7; Length 439;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTCTTCTGGATAC 18
Db      208 CCCCCCTCTGATTC 225
```

```
RESULT 43
US-09-925-065A-651102/c
; Sequence 651102, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 651102
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-651102
```

```
Query Match      82.2%; Score 14.8; DB 4; Length 547;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTCTTCTGGATAC 18
Db      260 CCCCCCTTCTGATTC 243
```

```
RESULT 44
US-10-437-963-48133
; Sequence 48133, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
```

```

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 48133
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50838C.1
; US-10-437-963-48133

```

```

Query Match      82.2%; Score 14.8; DB 7; Length 557;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Qy      1 CCCCCCTTCTTGATAC 18
          |||
Db      284 CCTCCCTCTCCGATAC 301

```

```

RESULT 45
; US-10-437-963-48132
; Sequence 48132, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 48132
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50837C.1
; US-10-437-963-48132

```

```

Query Match      82.2%; Score 14.8; DB 7; Length 564;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Qy      1 CCCCCCTTCTTGATAC 18
          |||
Db      147 CCTCCCTCTCCGATAC 164

```

```

RESULT 46
; US-09-925-065A-390692
; Sequence 390692, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

```

```

; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 390692
; LENGTH: 566
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-390692

```

```

Query Match      82.2%; Score 14.8; DB 4; Length 566;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Qy      1 CCCCCCTTCTTGATAC 18
          |||
Db      46 CCTGCCTCTCTGATAC 63

```

```

RESULT 47
; US-09-925-065A-736648
; Sequence 736648, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 736648
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-736648

```

```

Query Match      82.2%; Score 14.8; DB 4; Length 591;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Qy      1 CCCCCCTTCTTGATAC 18
          |||
Db      378 CCTCCCTCTCTGACAC 395

```

```

RESULT 48
; US-09-925-065A-736649
; Sequence 736649, Application US/09925065A

```



```
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 736649
/ LENGTH: 591
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-736649
```

```
Query Match      82.2% Score 14.8; DB 4; Length 591;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTTCTGTGATAC 18
          |||
Db      378 CTTCCCTTCTGTGACAC 395
```

```
RESULT 49
US-09-925-065A-814201
/ Sequence 814201, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 814201
/ LENGTH: 591
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-814201
```

```
Query Match      82.2% Score 14.8; DB 4; Length 591;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTTCTGTGATAC 18
          |||
Db      378 CTTCCCTTCTGTGACAC 395
```

```
RESULT 50
US-09-925-065A-814202
/ Sequence 814202, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 814202
/ LENGTH: 591
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-814202
```

```
Query Match      82.2% Score 14.8; DB 4; Length 591;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTTCTGTGATAC 18
          |||
Db      378 CTTCCCTTCTGTGACAC 395
```

```
RESULT 51
US-10-425-115-139469/c
/ Sequence 139469, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 139469
/ LENGTH: 604
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_58682C.1
US-10-425-115-139469
```

```
Query Match      82.2% Score 14.8; DB 8; Length 604;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTTCTGTGATAC 18
          |||
Db      182 CCCCCCTTCTGTGAGTTC 165
```

```
RESULT 52
US-09-925-065A-953026/c
/ Sequence 953026, Application US/09925065A
```

```
Publication No. US20050228172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIORITY FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIORITY FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIORITY FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIORITY FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIORITY FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIORITY FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 953026
LENGTH: 632
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-953026

Query Match      82.2%; Score 14.8; DB 4; Length 632;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTTCGATAC 18
      1 |||||
Db      26 CACCCCTCTTCGATCC 9

RESULT 53
US-09-925-065A-953455
Sequence 953455, Application US/09925065A
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIORITY FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIORITY FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIORITY FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIORITY FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIORITY FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIORITY FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 953455
LENGTH: 638
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-953455
```

```
Query Match      82.2%; Score 14.8; DB 4; Length 638;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTTCGATAC 18
      1 |||||
Db      613 CACCCCTCTTCGATCC 630
```

```
RESULT 54
US-10-027-632-21681/C
Sequence 21681, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIORITY FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIORITY FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIORITY FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIORITY FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIORITY FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIORITY FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIORITY FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIORITY FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21681
LENGTH: 679
TYPE: DNA
ORGANISM: Human
US-10-027-632-21681
```

```
Query Match      82.2%; Score 14.8; DB 5; Length 679;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 CCCCCCTCTTCGATAC 18
      1 |||||
Db      634 CCTCCCTCTTCGATAC 617
```

```
RESULT 55
US-10-027-632-21681/C
Sequence 21681, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIORITY FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIORITY FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIORITY FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIORITY FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIORITY FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIORITY FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIORITY FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIORITY FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21681
LENGTH: 679
TYPE: DNA
ORGANISM: Human
```

US-10-027-632-21681

Query Match 82.2%; Score 14.8; DB 6; Length 679;  
Best Local Similarity 88.9%; Pred. No. 1.3e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGGATAC 18  
DB 634 CCGCCCTCTCTGGATAC 617

RESULT 56

US-11-097-143-36419/c  
; Sequence 36419, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; FILE REFERENCE: CU000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; PRIOR FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36419  
; LENGTH: 1095  
; TYPE: DNA  
; ORGANISM: DROSOPHILA  
US-11-097-143-36419

Query Match 82.2%; Score 14.8; DB 10; Length 1095;  
Best Local Similarity 88.9%; Pred. No. 1.3e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGGATAC 18  
DB 188 CCGCCCTCTCTGGATAC 171

RESULT 57

US-10-450-763-28191  
; Sequence 28191, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; PRIOR FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23

; NUMBER OF SEQ ID NOS: 60736

; SOFTWARE: Custom

; SEQ ID NO 28191

; LENGTH: 1157

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIMILAR

; LOCATION: (975)..(1109)

; OTHER INFORMATION: 58% homologous to Homo sapiens MRJ, accession number

; OTHER INFORMATION: AB014888, Smith-Waterman Score=114.

US-10-450-763-28191

Query Match 82.2%; Score 14.8; DB 9; Length 1157;  
Best Local Similarity 88.9%; Pred. No. 1.3e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGGATAC 18  
DB 299 CCGCCCTCTCTGGATAC 316

RESULT 58

US-10-027-632-123726  
; Sequence 123726, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 123726  
; LENGTH: 1209  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-123726

Query Match 82.2%; Score 14.8; DB 5; Length 1209;  
Best Local Similarity 88.9%; Pred. No. 1.3e+03;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTCTCTGGATAC 18  
DB 878 CCGCCCTCTCTGGATAC 895

RESULT 59

US-10-027-632-123726  
; Sequence 123726, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129

```

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; APPLICATION NUMBER: US 60/198,576
; PRIOR FILING DATE: 2000-04-20
; APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 123726
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-123726
```

```

Query Match      82.2%; Score 14.8; DB 6; Length 1209;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 CCCCCCTCTTGATAC 18
    |||||
Db 878 CCGCTCTCTTGATAC 895
```

```

RESULT 60
US-09-925-065A-53648
; Sequence 53648, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 53648
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-53648
```

```

Query Match      82.2%; Score 14.8; DB 4; Length 1371;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 CCCCCCTCTTGATAC 18
    |||||
Db 624 CCGCTCTCTTGATAC 641
```

```

RESULT 61
US-09-925-065A-53649
```

```

; Sequence 53649, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 53649
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-53649
```

```

Query Match      82.2%; Score 14.8; DB 4; Length 1371;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 CCCCCCTCTTGATAC 18
    |||||
Db 624 CCGCTCTCTTGATAC 641
```

```

RESULT 62
US-09-925-065A-61730
; Sequence 61730, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 61730
; LENGTH: 1890
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-61730
```

```

Query Match      82.2%; Score 14.8; DB 4; Length 1890;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 CCCCCCTCTTGATAC 18
    |||||
Db 1645 CCGCTCTCTTGATAC 1662
```

```
RESULT 63
US-10-027-632-261924
; Sequence 261924, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 261924
; LENGTH: 1890
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-261924

Query Match      82.2%; Score 14.8; DB 5; Length 1890;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
; ORGANISM: Human
US-10-027-632-261925

Query Match      82.2%; Score 14.8; DB 5; Length 1890;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CCCCCCTCTCTGGATAC 18
Db      1645 CCTCTCTCTCTGGATAC 1662

RESULT 65
US-10-027-632-261924
; Sequence 261924, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 261924
; LENGTH: 1890
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-261924

Query Match      82.2%; Score 14.8; DB 6; Length 1890;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CCCCCCTCTCTGGATAC 18
Db      1645 CCTCTCTCTCTGGATAC 1662

RESULT 66
US-10-027-632-261925
; Sequence 261925, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
```

```

      PRIOR APPLICATION NUMBER: US 60/167,363
      PRIOR FILING DATE: 1999-11-23
      PRIOR APPLICATION NUMBER: US 60/156,358
      PRIOR FILING DATE: 1999-09-28
      PRIOR APPLICATION NUMBER: US 60/146,002
      PRIOR FILING DATE: 1999-08-09
      NUMBER OF SEQ ID NOS: 325720
      SOFTWARE: FastSeq for Windows Version 4.0
      SEQ ID NO 261925
      LENGTH: 1890
      TYPE: DNA
      ORGANISM: Human
US-10-027-632-261925

Query Match
Beat Local Similarity 82.2%; Score 14.8; DB 6; Length 1890;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCCCCCTTCTGTGATAC 18
    |||||
Db 1477 CCCCCCTCATCTGGGTAC 1662

RESULT 67
US-10-027-632-110327/c
; Sequence 110327, Application US/10027632
; Publication No. US20020198371a1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110327
; LENGTH: 2253
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-110327

Query Match
Beat Local Similarity 82.2%; Score 14.8; DB 5; Length 2253;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCCCCCTTCTGTGATAC 18
    |||||
Db 1477 CCCCCCTCATCTGGGTAC 1460

RESULT 68
US-10-027-632-110327/c
; Sequence 110327, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome

```

```

FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 110327
LENGTH: 2253
TYPE: DNA
ORGANISM: Human
US-10-027-632-110327

Query Match      82.2%   Score 14.8;   DB 6;   Length 2253;
Best Local Similarity 88.9%;   Pred. No. 1.3e+03;
Matches 16;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

1 CCCCCCTTCTGCATAC 18
|||||
Db 1477 CCCCCCTCATCTGGGTAC 1460

RESULT 69
US-10-739-930-5195/c
Sequence 5195, Application US/10739930
Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
FILE REFERENCE: 38-2153377B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 5195
LENGTH: 2297
TYPE: DNA
ORGANISM: Triticum aestivum
FEATURE:
OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER657_1
US-10-739-930-5195

Query Match      82.2%   Score 14.8;   DB 8;   Length 2297;
Best Local Similarity 88.9%;   Pred. No. 1.3e+03;
Matches 16;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

1 CCCCCCTTCTGCATAC 18
|||||
Db 2150 CACCCTCTTAGATAC 2133

RESULT 70
US-10-450-763-14809/c
Sequence 14809, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 14809
; LENGTH: 2433
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (418)..(1038)
; OTHER INFORMATION: 91% homologous to unidentified cloning vector 29kd protein
; OTHER INFORMATION: essential for the replication of mini F plasmid, accession number
; US-10-450-763-14809

```

```

Query Match      82.2%; Score 14.8; DB 9; Length 2433;
Best Local Similarity 88.9%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy      1 CCCCCCTCTTCTGGATAC 18
          ||| ||||| ||| |||
Db      168 CCCACCTCTTCTGGACAC 151

```

Search completed: May 5, 2006, 17:38:59  
 Job time : 500 secs

**THIS PAGE LEFT BLANK**



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using ew model

Run on: May 6, 2006, 19:05:31 ; Search time 385 Seconds  
(without alignments)  
190.404 Million cell updates/sec

Title: US-10-659-980A-5  
Perfect score: 18  
Sequence: 1 ccccccttcgatac 18

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9306428 segs, 2036268586 residues  
Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 70 summaries

Database :  
1: /SID55/ptodata/1/pubpna/US08\_NEW\_PUB.seq1:\*  
2: /SID55/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /SID55/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /SID55/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /SID55/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
6: /SID55/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
7: /SID55/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
8: /SID55/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
9: /SID55/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
10: /SID55/ptodata/1/pubpna/US10\_NEW\_PUB.seq1:\*  
11: /SID55/ptodata/1/pubpna/US10\_NEW\_PUB.seq3:\*  
12: /SID55/ptodata/1/pubpna/US10\_NEW\_PUB.seq4:\*  
13: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*  
14: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*  
15: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*  
16: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq4:\*  
17: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq4:\*  
18: /SID55/ptodata/1/pubpna/US11\_NEW\_PUB.seq4:\*  
19: /SID55/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15.4	85.6	104	US-10-310-914A-14889	Sequence 14889, A
2	15.4	85.6	623	US-10-301-480-429662	Sequence 429662,
3	15.4	85.6	623	US-10-301-480-1043071	Sequence 1043071,
4	15.4	85.6	623	US-09-925-065A-358412	Sequence 358412,
5	15.4	85.6	3260	US-09-925-065A-723345	Sequence 723345,
6	15.4	85.6	170189	US-11-112-908-50	Sequence 50, Appl
7	15.4	85.6	355211	US-10-330-773-242	Sequence 242, Appl
8	15.4	83.3	499	US-09-925-065A-166619	Sequence 166619,
9	15.4	83.3	522	US-10-301-480-258831	Sequence 258831,
10	15.4	83.3	522	US-10-301-480-872240	Sequence 872240,
11	15.4	83.3	2776	US-11-072-512-1147	Sequence 1147, Ap
12	14.8	82.2	494	US-10-301-480-466521	Sequence 466521,
13	14.8	82.2	494	US-10-301-480-466522	Sequence 466522,
14	14.8	82.2	494	US-10-301-480-1079930	Sequence 1079930,

## ALIGNMENTS

C 15	14.8	82.2	494	12	US-10-301-480-1079931	Sequence 1079931,
C 16	14.8	82.2	547	7	US-09-925-065A-651102	Sequence 651102,
C 17	14.8	82.2	564	12	US-10-301-480-458989	Sequence 458989,
C 18	14.8	82.2	564	12	US-10-301-480-1072398	Sequence 1072398,
C 19	14.8	82.2	566	7	US-09-925-065A-390632	Sequence 390632,
C 20	14.8	82.2	591	7	US-09-925-065A-736648	Sequence 736648,
C 21	14.8	82.2	591	7	US-09-925-065A-736649	Sequence 736649,
C 22	14.8	82.2	591	7	US-09-925-065A-814201	Sequence 814201,
C 23	14.8	82.2	591	7	US-09-925-065A-814202	Sequence 814202,
C 24	14.8	82.2	632	7	US-09-925-065A-953026	Sequence 953026,
C 25	14.8	82.2	632	7	US-09-925-065A-953026	Sequence 953026,
C 26	14.8	82.2	744	12	US-10-301-480-546939	Sequence 546939,
C 27	14.8	82.2	744	12	US-10-301-480-1160408	Sequence 1160408,
C 28	14.8	82.2	993	12	US-10-301-480-536495	Sequence 536495,
C 29	14.8	82.2	993	12	US-10-301-480-552560	Sequence 552560,
C 30	14.8	82.2	993	12	US-10-301-480-1149904	Sequence 1149904,
C 31	14.8	82.2	1371	7	US-10-301-480-1165969	Sequence 1165969,
C 32	14.8	82.2	1371	7	US-09-925-065A-53648	Sequence 53648, A
C 33	14.8	82.2	1371	7	US-10-301-480-154886	Sequence 154886,
C 34	14.8	82.2	1371	11	US-10-301-480-154887	Sequence 154887,
C 35	14.8	82.2	1371	11	US-10-301-480-768235	Sequence 768235,
C 36	14.8	82.2	1371	12	US-10-301-480-768236	Sequence 768236,
C 37	14.8	82.2	1371	12	US-10-301-480-768236	Sequence 768236,
C 38	14.8	82.2	1890	7	US-09-925-065A-617130	Sequence 617130, A
C 39	14.8	82.2	1890	11	US-10-301-480-96741	Sequence 96741, A
C 40	14.8	82.2	1890	11	US-10-301-480-152968	Sequence 152968,
C 41	14.8	82.2	1890	11	US-10-301-480-152968	Sequence 152968,
C 42	14.8	82.2	1890	12	US-10-301-480-710150	Sequence 710150,
C 43	14.8	82.2	1890	12	US-10-301-480-710151	Sequence 710151,
C 44	14.8	82.2	1890	12	US-10-301-480-710151	Sequence 710151,
C 45	14.8	82.2	2671	10	US-10-750-185-42813	Sequence 42813, A
C 46	14.8	82.2	2671	10	US-10-750-185-42813	Sequence 42813, A
C 47	14.8	82.2	3216	10	US-10-750-185-28164	Sequence 28164, A
C 48	14.8	82.2	3216	10	US-10-750-185-28164	Sequence 28164, A
C 49	14.8	82.2	3422	10	US-10-750-185-47185	Sequence 47185, A
C 50	14.8	82.2	3422	10	US-10-750-185-47185	Sequence 47185, A
C 51	14.8	82.2	66648	11	US-10-330-773-695	Sequence 695, Appl
C 52	14.8	82.2	109147	11	US-10-330-773-695	Sequence 695, Appl
C 53	14.8	82.2	163182	17	US-11-121-086-66	Sequence 66, Appl
C 54	14.8	82.2	198977	18	US-11-114-798-54	Sequence 54, Appl
C 55	14.8	82.2	559662	8	US-09-824-575A-3	Sequence 3, Appl1
C 56	14.4	80.0	201	10	US-10-995-561-49194	Sequence 49194, A
C 57	14.4	80.0	201	17	US-11-124-368A-13541	Sequence 13541, A
C 58	14.4	80.0	201	17	US-11-124-368A-13597	Sequence 13597, A
C 59	14.4	80.0	201	17	US-11-124-368A-20109	Sequence 20109, A
C 60	14.4	80.0	421	7	US-09-925-065A-478927	Sequence 478927,
C 61	14.4	80.0	421	12	US-10-301-480-357026	Sequence 357026,
C 62	14.4	80.0	479	12	US-10-301-480-970425	Sequence 970425,
C 63	14.4	80.0	491	12	US-10-301-480-564341	Sequence 564341,
C 64	14.4	80.0	491	12	US-10-301-480-1177750	Sequence 1177750,
C 65	14.4	80.0	523	7	US-09-925-065A-259422	Sequence 259422,
C 66	14.4	80.0	523	7	US-09-925-065A-259423	Sequence 259423,
C 67	14.4	80.0	529	7	US-09-925-065A-244698	Sequence 244698,
C 68	14.4	80.0	531	12	US-10-301-480-337948	Sequence 337948,
C 69	14.4	80.0	531	12	US-10-301-480-951357	Sequence 951357,
C 70	14.4	80.0	545	7	US-09-925-065A-791113	Sequence 791113,

## RESULT 1

US-10-310-914A-14889  
; Sequence 14889, Application US/10310914A  
; Publication No. US20060003322A1  
; GENERAL INFORMATION:  
; APPLICANT: Bentwich, Isaac  
; APPLICANT: Shiner, Kyuzat  
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and  
; FILE REFERENCE: 06087.0200.CPUS01  
; CURRENT APPLICATION NUMBER: US/10/310,914A  
; CURRENT FILING DATE: 2002-12-06

```
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 14889
; LENGTH: 104
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-14889
```

```
Query Match      85.6%; Score 15.4; DB 10; Length 104;
Best Local Similarity 64.7%; Pred. No. 1.6e+02;
Matches 11; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2 CCCCCCTTCTGTGATAC 18
Db      82 CCCCCCUCUGGAAUUC 98
```

```
RESULT 2
US-10-301-480-429662/C
; Sequence 429662, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO: 429662
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-429662
```

```
Query Match      85.6%; Score 15.4; DB 12; Length 623;
Best Local Similarity 94.1%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTTCTGTGATA 17
Db      283 CCCCCCTTATGAGATA 267
```

```
RESULT 3
US-10-301-480-1043071/C
; Sequence 1043071, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO: 1043071
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-1043071
```

```
Query Match      85.6%; Score 15.4; DB 12; Length 623;
```

```
Best Local Similarity 94.1%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTTCTGTGATA 17
Db      283 CCCCCCTTATGAGATA 267
```

```
RESULT 4
US-09-925-065A-358412/C
; Sequence 358412, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO: 358412
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-358412
```

```
Query Match      85.6%; Score 15.4; DB 7; Length 630;
Best Local Similarity 94.1%; Pred. No. 2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 CCCCCCTTCTGTGATA 17
Db      290 CCCCCCTTATGAGATA 274
```

```
RESULT 5
US-09-925-065A-723345
; Sequence 723345, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO: 723345
; LENGTH: 3263
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-925-065A-723345

Query Match 85.6%; Score 15.4; DB 7; Length 3263;  
Best Local Similarity 94.1%; Pred. No. 2.4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCCTTCTGTGATA 17  
|||  
Db 268 CCCCCCTTCTGTGATA 284

RESULT 6  
US-11-112-908-50/c

; Sequence 50, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; PRIOR FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 50  
; LENGTH: 170189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-50

Query Match 85.6%; Score 15.4; DB 17; Length 170189;  
Best Local Similarity 94.1%; Pred. No. 3.7e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCCCCTTCTGTGATAC 18  
|||  
Db 29188 CCCCCCTTCTGTGATAC 29172

RESULT 7  
US-10-330-773-242/c

; Sequence 242, Application US/10330773  
; Publication No. US20060040262A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: 529452001300  
; CURRENT APPLICATION NUMBER: US/10/330,773  
; CURRENT FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 991  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 242  
; LENGTH: 355211  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(355211)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-330-773-242

Query Match 85.6%; Score 15.4; DB 11; Length 355211;  
Best Local Similarity 94.1%; Pred. No. 4e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCCCCTTCTGTGATAC 18  
|||  
Db 102906 CACCCCTTCTGTGATAC 102890

RESULT 8  
US-09-925-065A-166619/c

; Sequence 166619, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 166619  
; LENGTH: 499  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-166619

Query Match 83.3%; Score 15; DB 7; Length 499;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCCCCTTCTGTGAT 16  
|||  
Db 182 CCCCCCTTCTGTGAT 168

RESULT 9  
US-10-301-480-258831/c

; Sequence 258831, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 258831  
; LENGTH: 522  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-258831

Query Match 83.3%; Score 15; DB 12; Length 522;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCCCCTTCTGTGAT 16  
|||

Db 205 CCCCTCTTCTGGAT 191

RESULT 10  
US-10-301-480-872240/c  
; Sequence 872240, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE REFERENCE: 108827.137

; CURRENT APPLICATION NUMBER: US/10/301,480

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; PRIOR FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 872240

; LENGTH: 522

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-301-480-872240

Query Match 83.3%; Score 15; DB 12; Length 522;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCCTCTTCTGGAT 16

Db 205 CCCCTCTTCTGGAT 191

RESULT 11  
US-11-072-512-1147/c

; Sequence 1147, Application US/11072512

; Publication No. US20060029945A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHICO

; APPLICANT: YOSHITAKA, TSUTOMU

; APPLICANT: OTSUKA, MOTOKYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: Novel full length cDNA

; FILE REFERENCE: 084335-0191

; CURRENT APPLICATION NUMBER: US/11/072,512

; CURRENT FILING DATE: 2005-03-07

; PRIOR APPLICATION NUMBER: US 60/350,978

; PRIOR FILING DATE: 2002-01-25

; PRIOR APPLICATION NUMBER: JP 2001-379298

; PRIOR FILING DATE: 2001-11-05

; NUMBER OF SEQ ID NOS: 4096

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1147

; LENGTH: 2776

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-072-512-1147

Query Match 83.3%; Score 15; DB 18; Length 2776;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CCCCTCTTCTGGATAC 18  
Db 2173 CCCCTCTTCTGGATAC 2159

RESULT 12  
US-10-301-480-466521/c

; Sequence 466521, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE REFERENCE: 108827.137

; CURRENT APPLICATION NUMBER: US/10/301,480

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; PRIOR FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 466521

; LENGTH: 494

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-301-480-466521

Query Match 82.2%; Score 14.8; DB 12; Length 494;  
Best Local Similarity 88.9%; Pred. No. 4.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCCTCTTCTGGATAC 18  
Db 39 CCCCTCTTCTGGATAC 22

RESULT 13

US-10-301-480-466522/c

; Sequence 466522, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE REFERENCE: 108827.137

; CURRENT APPLICATION NUMBER: US/10/301,480

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; PRIOR FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 466522

; LENGTH: 494

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-301-480-466522

Query Match 82.2%; Score 14.8; DB 12; Length 494;  
Best Local Similarity 88.9%; Pred. No. 4.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCCTCTTCTGGATAC 18  
Db 39 CCCCTCTTCTGGATAC 22

```
RESULT 14
US-10-301-480-1079930/c
; Sequence 1079930, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1079930
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1079930

Query Match      82.2%; Score 14.8; DB 12; Length 494;
Best Local Similarity 88.9%; Pred. No. 4.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGATAC 18
    |||||
Db 39 CCCCCCTTCTTGATAC 22

RESULT 15
US-10-301-480-1079931/c
; Sequence 1079931, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1079931
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1079931

Query Match      82.2%; Score 14.8; DB 12; Length 494;
Best Local Similarity 88.9%; Pred. No. 4.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGATAC 18
    |||||
Db 39 CCCCCCTTCTTGATAC 22

RESULT 16
US-09-925-065A-651102/c
; Sequence 651102, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
```

```
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 651102
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-651102

Query Match      82.2%; Score 14.8; DB 7; Length 547;
Best Local Similarity 88.9%; Pred. No. 4.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGATAC 18
    |||||
Db 260 CCCCCCTTCTTGATAC 243

RESULT 17
US-10-301-480-458989
; Sequence 458989, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458989
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-458989

Query Match      82.2%; Score 14.8; DB 12; Length 564;
Best Local Similarity 88.9%; Pred. No. 4.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTTGATAC 18
    |||||
Db 44 CCCCCCTTCTTGATAC 61

RESULT 18
US-10-301-480-1072398
; Sequence 1072398, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
```

```

1 PRIOR APPLICATION NUMBER: US 10/215,598
2
3 PRIOR FILING DATE: 2002-08-09
4
5 PRIOR APPLICATION NUMBER: US 60/311,695
6
7 PRIOR FILING DATE: 2001-08-10
8
9 NUMBER OF SEQ ID NOS: 122618
10
11 SOFTWARE: FASTSEQ for Windows Version 4.0
12
13 SEQ ID NO 1072398
14
15 LENGTH: 564
16
17 TYPE: DNA
18
19 ORGANISM: Homo sapien
20
21 US-10-301480-1072398

```

Query Match	82.2%	Score 14.8;	DB 12;	Length 564;
Best Local Similarity	88.9%	Pred. No. 4.2e+02;		
Matches 16;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

OY		CCCCCTTCTGATAC	18
Db	44	CCTGCCTCTTGATAC	61

RESULT 19  
US-09-925-065A-390692

```

Sequence 390692. Application US/09925065A
Publication NO. US2004018104A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,036
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 390692
LENGTH: 566
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-390692

```

Query Match	82.2%	Score 14.8;	DB 7;	Length 566;
Best Local Similarity	88.9%	Pred. No. 4.2e+02;		
Matches 16; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

Oy		1	CCCCCTTCTGATAC	18
Dd		46	CCTGCCTTTCTGGATAC	63

RESULT 20  
US-09-925-065A-736648  
Sequence 736648, Application US/09925065A  
Publication No. US20040181048A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single  
Nucleotide Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.15  
CURRENT APPLICATION NUMBER: US/09/925,065A  
CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147

```

? PRIOR FILING DATE: 2000-11-20
? PRIOR APPLICATION NUMBER: US 60/250, 092
? PRIOR FILING DATE: 2000-11-30
? PRIOR APPLICATION NUMBER: US 60/261, 766
? PRIOR FILING DATE: 2001-01-16
? PRIOR APPLICATION NUMBER: US 60/289, 846
? PRIOR FILING DATE: 2001-05-09
? NUMBER OF SEQ ID NOS: 957086
? SOFTWARE: FASTSEQ for Windows Version 4.0.0
? SEQ ID NO 736648
? LENGTH: 591
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-925-065A--736648

```

Query Match	82.2%	Score 14.8	DB 7	Length 591
Best Local Similarity	88.9%	Pred. No. 4.2e+02		
Matches 16	Conservative 0	Mismatches 2	Indels 0	Gaps 0

```

Qy      1  CCCCCCTTCTGTGATAC  18
          |||||
Db      378 CCTCCCTCTTCTGTGACAC  395

```

```

RESULT 21
US-09-925-065A-736649
; Sequence 736649, Application US/09925065A
; Publication No. US20040181084A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108877.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/269,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 736649
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-736649

```

Query Match	82.2%	Score 14.8	DB 7	Length 591
Best Local Similarity	88.9%	Pred. No. 4.2e+02		
Matches 16; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

```

QY      1  CCCCCCTTCTGATAC  18
          |||||
Db      378 CCTCCCTCTTCTGACAC  395

```

```

RESULT 22
US-09-925-065A-814201
; Sequence 814201, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08

```

```

; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 814201
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-814201

```

Query Match	82.2%	Score 14.8	DB 7	Length 591
Best Local Similarity	88.9%	Pred. No. 4,2e+02		
Matches	16	Conservative	0	Mismatches 2
				Indels 0
				Gaps 0
OY	1	CCCCCTCTTGCAATAC	18	
db	378	CTCTCCCTCTTGACAC	395	

```

RESULT 23
US-09-925-065A--814202
; Sequence 814202, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single

```

```

1  TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
2
3  FILE REFERENCE: 108827.135
4
5  CURRENT APPLICATION NUMBER: US/09/925,065A
6
7  CURRENT FILING DATE: 2001-08-08
8
9  PRIOR APPLICATION NUMBER: US 60/243,096
10
11 PRIOR FILING DATE: 2000-10-24
12
13 PRIOR APPLICATION NUMBER: US 60/252,147
14
15 PRIOR FILING DATE: 2000-11-20

```

```

?
? PRIOR APPLICATION NUMBER: US 60/250,052
?
? PRIOR FILING DATE: 2000-11-30
?
? PRIOR APPLICATION NUMBER: US 60/261,766
?
? PRIOR FILING DATE: 2001-01-16
?
? PRIOR APPLICATION NUMBER: US 60/289,846
?
? PRIOR FILING DATE: 2001-05-09
?
? NUMBER OF SEQ ID NOS: 957086
?
? SOFTWARE: FastSeq for Windows Version 4.0
?
? SEQ ID NO. 814202

```

```

; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-814202

Query Match      82.2%   Score 14.8; DB 7; Length 591;
Best Local Similarity 88.9%   Pred. No. 4.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

```

Qy      1  CCCCCCTTTGGATAC  18
        |||||
Db      378  CCTCCCTTTGGACAC  395

```

```

: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single
: TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

```

```

? FILE REFERENCE: 108827.135
? CURRENT APPLICATION NUMBER: US/09/925,065A
? CURRENT FILING DATE: 2001-08-08
? PRIOR APPLICATION NUMBER: US 60/243,096
? PRIOR FILING DATE: 2000-10-24
? PRIOR APPLICATION NUMBER: US 60/252,147
? PRIOR FILING DATE: 2000-11-20
? PRIOR APPLICATION NUMBER: US 60/250,092
? PRIOR FILING DATE: 2000-11-30
? PRIOR APPLICATION NUMBER: US 60/261,766
? PRIOR FILING DATE: 2001-01-16
? PRIOR APPLICATION NUMBER: US 60/289,846
? PRIOR FILING DATE: 2001-05-09
? NUMBER OF SEQ. ID NOS: 957086
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 953026
? LENGTH: 632
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-925-065A-953026

```

Query Match	82.2%;	Score 14.9;	DB 7;	Length 632;
Best Local Similarity	88.9%;	Pred. No. 4.2e+02;		
Matches	16;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;
Qy	1	CCCCCTCTTGTGATAC	18	
Db	26	CACCCCTCTTGTGATCC	9	

RESULT 25  
US-09-925-065A-953455  
; Sequence 953455, Application US/09925065A  
; Publication No. US2004018104B1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single

TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
 FILE REFERENCE: 108827.135  
 CURRENT APPLICATION NUMBER: US/09/925,065A  
 CURRENT FILING DATE: 2001-08-08  
 PRIOR APPLICATION NUMBER: US 60/243,036  
 PRIOR FILING DATE: 2000-10-24  
 PRIOR APPLICATION NUMBER: US 60/252,147  
 PRIOR FILING DATE: 2000-11-20

```

?
? PRIOR APPLICATION NUMBER: US 60/250,022
?
? PRIOR FILING DATE: 2000-11-30
?
? PRIOR APPLICATION NUMBER: US 60/261,766
?
? PRIOR FILING DATE: 2001-01-16
?
? PRIOR APPLICATION NUMBER: US 60/289,846
?
? PRIOR FILING DATE: 2001-05-09
?
? NUMBER OF SEQ ID NOS: 957086
?
? SOFTWARE: FastSeq for Windows Version 4.0
?
? SEQ ID NO 957455

```

```

?      LENGTH: 638
?      TYPE: DNA
?      ORGANISM: Homo sapiens
US-09-925-065A-953455

Query Match      Score 14.8;  DB 7;      length 638;
Best Local Similarity 88.9%;      Pred No. 4.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0

```

```

QY      1  CCCCCCTTCTGATAC  18
          |||||
Db      613  CCCCCCTTCTGATAC  630

```

US-10-301-480-54699/c  
; Sequence 54699, Application US/10301480  
; Publication No. US20060057564v1  
; GENERAL INFORMATION:

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
3  TITLE OF INVENTION: in the Human Genome
4  FILE REFERENCE: 108827.137
5  CURRENT APPLICATION NUMBER: US/10/301,480
6  CURRENT FILING DATE: 2002-11-21
7  PRIOR APPLICATION NUMBER: US 10/215,598
8  PRIOR FILING DATE: 2002-08-09
9  PRIOR APPLICATION NUMBER: US 60/311,695
10 PRIOR FILING DATE: 2001-08-10
11 NUMBER OF SEQ ID NOS: 1226818
12 SOFTWARE: FastSeq for Windows Version 4.0
13 SEQ ID NO 546999
14 LENGTH: 744
15 TYPE: DNA
16 ORGANISM: Homo sapien
17 US-10-301-480-546999

```

Query Match	82.2%	Score 14.8;	DB 12;	Length 744;
Best Local Similarity	88.9%	Pred. No. 4.3e+02;		
Matches 16;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY      1 CCCCCCTTCTGATAC 18
          | ||||| ||| |||||
Db      111 CTCCTCTTCAGATAC 94
```

```

RESULT 27
US-10-301-480-1160408/c
? Sequence 1160408, Application US/10301480
? Publication No. US20060057564A1
? GENERAL INFORMATION:
? APPLICANT: Wang, David G.
? TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
? TITLE OF INVENTION: in the Human Genome
? FILE REFERENCE: 108827.137
? CURRENT APPLICATION NUMBER: US/10/301,480
? CURRENT FILING DATE: 2002-11-21
? PRIOR APPLICATION NUMBER: US 10/215,598
? PRIOR FILING DATE: 2002-08-09
? PRIOR APPLICATION NUMBER: US 60/311,695
? PRIOR FILING DATE: 2001-08-10
? NUMBER OF SEQ ID NOS: 1226818
? SOFTWARE: SeqSeq for Windows Version 4.0
? SEQ ID NO 1160408
?
? LENGTH: 744
?
? TYPE: DNA
? ORGANISM: Homo sapien
US-10-301-480-1160408

```

Query Match	82.2%	Score 14.8;	DB 12;	Length 744;
Best Local Similarity	88.9%	Pred. No. 4.3e+02;		
Matches 16;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	1	CCCCCTCTTGATAC	18
Db	111	CTCCCTCTTCAGATAC	94

```

RESULT 28
US-10-301-480-536495
: Sequence 536495, Application US/10301480
: Publication No. US20060057564A1
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
: TITLE OF INVENTION: in the Human Genome
: PTE REFERENCE: 108927.137
: CURRENT APPLICATION NUMBER: US/10/301,480
: CURRENT FILING DATE: 2002-11-21
: PRIOR APPLICATION NUMBER: US 10/215,598
: PRIOR FILING DATE: 2002-08-09
: PRIOR APPLICATION NUMBER: US 60/311,595

```

```

? PRIOR FILING DATE: 2001-08-10
? NUMBER OF SEQ ID NOS: 1226818
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 536495
? LENGTH: 993
? TYPE: DNA
? ORGANISM: Homo sapien
US-10-301-480-536495

```

Query Match	82.2%	Score 14.8;	DB 12;	Length 993;
Best Local Similarity	88.9%	Pred. No. 4.5e+02;		
Matches 16;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

OY		1	CCCCCTTCTGATAC	18
D6		245	CCACCTTCTTGATAC	262

```

RESULT 29
US-10-301-480-552560
; Sequence 552560, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 552560
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-552560

```

Query Match	82.2%	Score 14.8;	DB 12;	Length 993;
Best Local Similarity	88.9%	Pred. No. 4.5e+02;		
Matches 16; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY		1	CCCCCTTCTGATAC	18
Db		829	CCCCCTTCTTGACAC	846

```

RESULT 30
US-10-301-480-1149904
; Sequence 1149904, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1149904
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1149904

```



Query Match 82.2%; Score 14.8; DB 12; Length 993;  
Best Local Similarity 88.9%; Pred. No. 4.5e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGTGATAC 18  
DB 245 CCACCTTCTGTGATAC 262

RESULT 31  
US-10-301-480-1165969  
; Sequence 1165969, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1165969  
; LENGTH: 993  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-1165969

Query Match 82.2%; Score 14.8; DB 12; Length 993;  
Best Local Similarity 88.9%; Pred. No. 4.5e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGTGATAC 18  
DB 829 CCCCCCTTCTGTGATAC 846

RESULT 32  
US-09-925-065A-53648  
; Sequence 53648, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 53648  
; LENGTH: 1371  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-53648

Query Match 82.2%; Score 14.8; DB 7; Length 1371;  
Best Local Similarity 88.9%; Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGTGATAC 18  
DB 624 CCCCCCTTCTGTGATAC 641

RESULT 33  
US-09-925-065A-53649  
; Sequence 53649, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 53649  
; LENGTH: 1371  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-53649

Query Match 82.2%; Score 14.8; DB 7; Length 1371;  
Best Local Similarity 88.9%; Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGTGATAC 18  
DB 624 CCCCCCTTCTGTGATAC 641

RESULT 34  
US-10-301-480-154886  
; Sequence 154886, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 154886  
; LENGTH: 1371  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-154886

Query Match 82.2%; Score 14.8; DB 11; Length 1371;  
Best Local Similarity 88.9%; Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTCTGTGATAC 18  
DB 624 CCCCCCTTCTGTGATAC 641

DB 624 CCCCTCTCTTGATGC 641

```
RESULT 35
US-10-301-480-15487
; Sequence 15487, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15487
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-15487
```

Query Match 82.2%; Score 14.8; DB 11; Length 1371;  
Best Local Similarity 88.9%; Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCTCTCTTGATGC 18  
DB 624 CCCCTCTCTTGATGC 641

```
RESULT 36
US-10-301-480-768295
; Sequence 768295, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 768295
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-768295
```

Query Match 82.2%; Score 14.8; DB 12; Length 1371;  
Best Local Similarity 88.9%; Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCTCTCTTGATGC 18  
DB 624 CCCCTCTCTTGATGC 641

```
RESULT 37
US-10-301-480-768296
; Sequence 768296, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 768296
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-768296
```

Query Match 82.2%; Score 14.8; DB 12; Length 1371;  
Best Local Similarity 88.9%; Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCTCTCTTGATGC 18  
DB 624 CCCCTCTCTTGATGC 641

```
RESULT 38
US-09-925-065A-61730
; Sequence 61730, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61730
; LENGTH: 1890
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-61730
```

Query Match 82.2%; Score 14.8; DB 7; Length 1890;  
Best Local Similarity 88.9%; Pred. No. 4.8e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCTCTCTTGATGC 18  
DB 1645 CCCTCTCTTGATGC 1662

```
RESULT 39
US-10-301-480-96741
; Sequence 96741, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
```

```

;
;   CURRENT APPLICATION NUMBER: US/10/301,480
;   CURRENT FILING DATE: 2002-11-21
;   PRIOR APPLICATION NUMBER: US 10/215,598
;   PRIOR FILING DATE: 2002-08-09
;   PRIOR APPLICATION NUMBER: US 60/311,695
;   PRIOR FILING DATE: 2001-08-10
;   NUMBER OF SEQ ID NOS: 1226818
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 96741
;   LENGTH: 1890
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   US-10-301-480-96741

Query Match      82.2%; Score 14.8; DB 11; Length 1890;
Best Local Similarity 88.9%; Pred. No. 4.8e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CCCCCCTCTCTGGATAC 18
Db      1645 CCTCTCTCTCTGGGATAC 1662

RESULT 40
US-10-301-480-96742
; Sequence 96742, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
;   FILE REFERENCE: 108827.137
;   CURRENT APPLICATION NUMBER: US/10/301,480
;   CURRENT FILING DATE: 2002-11-21
;   PRIOR APPLICATION NUMBER: US 10/215,598
;   PRIOR FILING DATE: 2002-08-09
;   PRIOR APPLICATION NUMBER: US 60/311,695
;   PRIOR FILING DATE: 2001-08-10
;   NUMBER OF SEQ ID NOS: 1226818
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 96742
;   LENGTH: 1890
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   US-10-301-480-96742

Query Match      82.2%; Score 14.8; DB 11; Length 1890;
Best Local Similarity 88.9%; Pred. No. 4.8e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CCCCCCTCTCTGGATAC 18
Db      1645 CCTCTCTCTCTGGGATAC 1662

RESULT 41
US-10-301-480-162968
; Sequence 162968, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
;   FILE REFERENCE: 108827.137
;   CURRENT APPLICATION NUMBER: US/10/301,480
;   CURRENT FILING DATE: 2002-11-21
;   PRIOR APPLICATION NUMBER: US 10/215,598
;   PRIOR FILING DATE: 2002-08-09
;   PRIOR APPLICATION NUMBER: US 60/311,695
;   PRIOR FILING DATE: 2001-08-10
;   NUMBER OF SEQ ID NOS: 1226818
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 162968
```

```

;   LENGTH: 1890
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   US-10-301-480-162968

Query Match      82.2%; Score 14.8; DB 11; Length 1890;
Best Local Similarity 88.9%; Pred. No. 4.8e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CCCCCCTCTCTGGATAC 18
Db      1645 CCTCTCTCTCTGGGATAC 1662

RESULT 42
US-10-301-480-710150
; Sequence 710150, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
;   FILE REFERENCE: 108827.137
;   CURRENT APPLICATION NUMBER: US/10/301,480
;   CURRENT FILING DATE: 2002-11-21
;   PRIOR APPLICATION NUMBER: US 10/215,598
;   PRIOR FILING DATE: 2002-08-09
;   PRIOR APPLICATION NUMBER: US 60/311,695
;   PRIOR FILING DATE: 2001-08-10
;   NUMBER OF SEQ ID NOS: 1226818
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 710150
;   LENGTH: 1890
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   US-10-301-480-710150

Query Match      82.2%; Score 14.8; DB 12; Length 1890;
Best Local Similarity 88.9%; Pred. No. 4.8e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CCCCCCTCTCTGGATAC 18
Db      1645 CCTCTCTCTCTGGGATAC 1662

RESULT 43
US-10-301-480-710151
; Sequence 710151, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
;   FILE REFERENCE: 108827.137
;   CURRENT APPLICATION NUMBER: US/10/301,480
;   CURRENT FILING DATE: 2002-11-21
;   PRIOR APPLICATION NUMBER: US 10/215,598
;   PRIOR FILING DATE: 2002-08-09
;   PRIOR APPLICATION NUMBER: US 60/311,695
;   PRIOR FILING DATE: 2001-08-10
;   NUMBER OF SEQ ID NOS: 1226818
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 710151
;   LENGTH: 1890
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   US-10-301-480-710151

Query Match      82.2%; Score 14.8; DB 12; Length 1890;
Best Local Similarity 88.9%; Pred. No. 4.8e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 CCCCCCTTTTGATAC 18  
 |||||  
 DB 1645 CCCTCTTTTGGCTAC 1662

RESULT 44  
 US-10-301-480-776377  
 ; Sequence 776377, Application US/10301480  
 ; Publication No. US20060057564A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
 ; FILE REFERENCE: 108827.137  
 ; CURRENT APPLICATION NUMBER: US/10/301,480  
 ; PRIOR FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US 10/215,598  
 ; PRIOR FILING DATE: 2002-08-09  
 ; PRIOR APPLICATION NUMBER: US 60/311,695  
 ; NUMBER OF SEQ ID NOS: 122618  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 776377  
 ; LENGTH: 1890  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 US-10-301-480-776377

Query Match 82.2%; Score 14.8; DB 12; Length 1890;  
 Best Local Similarity 88.9%; Pred. No. 4.8e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTTTGATAC 18  
 |||||  
 DB 1645 CCCTCTTTTGGCTAC 1662

RESULT 45  
 US-10-750-185-42813/c  
 ; Sequence 42813, Application US/10750185  
 ; Publication No. US20050260603A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MMI GENOMICS, INC.  
 ; APPLICANT: DENISE, Sue K.  
 ; APPLICANT: KERR, Richard  
 ; APPLICANT: ROSENFELD, David  
 ; APPLICANT: HOLM, Tom  
 ; APPLICANT: BATES, Stephen  
 ; APPLICANT: FANTIN, Dennis  
 ; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
 ; FILE REFERENCE: MM11100-2  
 ; CURRENT APPLICATION NUMBER: US/10/750,185  
 ; CURRENT FILING DATE: 2003-12-31  
 ; PRIOR APPLICATION NUMBER: US 60/437,482  
 ; PRIOR FILING DATE: 2002-12-31  
 ; NUMBER OF SEQ ID NOS: 64922  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 42813  
 ; LENGTH: 2671  
 ; TYPE: DNA  
 ; ORGANISM: Bovine  
 US-10-750-185-42813

Query Match 82.2%; Score 14.8; DB 10; Length 2671;  
 Best Local Similarity 88.9%; Pred. No. 5e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTTTGATAC 18  
 |||||  
 DB 361 CCCCCCTTTTGCTAC 344

RESULT 46

US-10-750-623-42813/c  
 ; Sequence 42813, Application US/10750623  
 ; Publication No. US20050287531A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MMI GENOMICS, INC.  
 ; APPLICANT: DENISE, Sue K.  
 ; APPLICANT: KERR, Richard  
 ; APPLICANT: ROSENFELD, David  
 ; APPLICANT: HOLM, Tom  
 ; APPLICANT: BATES, Stephen  
 ; APPLICANT: FANTIN, Dennis  
 ; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS  
 ; FILE REFERENCE: MM11100-1  
 ; CURRENT APPLICATION NUMBER: US/10/750,623  
 ; CURRENT FILING DATE: 2003-12-31  
 ; PRIOR APPLICATION NUMBER: US 60/437,482  
 ; PRIOR FILING DATE: 2002-12-31  
 ; NUMBER OF SEQ ID NOS: 64922  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 42813  
 ; LENGTH: 2671  
 ; TYPE: DNA  
 ; ORGANISM: Bovine  
 US-10-750-623-42813

Query Match 82.2%; Score 14.8; DB 10; Length 2671;  
 Best Local Similarity 88.9%; Pred. No. 5e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTTTGATAC 18  
 |||||  
 DB 361 CCCCCCTTTTGCTAC 344

RESULT 47  
 US-10-750-185-28164/c  
 ; Sequence 28164, Application US/10750185  
 ; Publication No. US20050260603A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MMI GENOMICS, INC.  
 ; APPLICANT: DENISE, Sue K.  
 ; APPLICANT: KERR, Richard  
 ; APPLICANT: ROSENFELD, David  
 ; APPLICANT: HOLM, Tom  
 ; APPLICANT: BATES, Stephen  
 ; APPLICANT: FANTIN, Dennis  
 ; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
 ; FILE REFERENCE: MM11100-2  
 ; CURRENT APPLICATION NUMBER: US/10/750,185  
 ; CURRENT FILING DATE: 2003-12-31  
 ; PRIOR APPLICATION NUMBER: US 60/437,482  
 ; PRIOR FILING DATE: 2002-12-31  
 ; NUMBER OF SEQ ID NOS: 64922  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 28164  
 ; LENGTH: 3216  
 ; TYPE: DNA  
 ; ORGANISM: Bovine  
 US-10-750-185-28164

Query Match 82.2%; Score 14.8; DB 10; Length 3216;  
 Best Local Similarity 88.9%; Pred. No. 5.1e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCCCCTTTTGATAC 18  
 |||||  
 DB 106 CCACCTTACTGATAC 89

RESULT 48  
 US-10-750-623-28164/c  
 ; Sequence 28164, Application US/10750623  
 ; Publication No. US20050287531A1

```

; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28164
; LENGTH: 3216
; TYPE: DNA
; ORGANISM: Bovine 19866880735400
US-10-750-623-28164

Query Match      82.2%  Score 14.8; DB 10; Length 3216;
Best Local Similarity 88.9%  Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1  CCCCCCTCTCTGGATAC 18
DB      106  CCCACCTCTCTGGATAC 89

RESULT 49
US-10-750-185-47185
; Sequence 47185, Application US/10750185
; Publication No. US2005026603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47185
; LENGTH: 3422
; TYPE: DNA
; ORGANISM: Bovine 19866881073042
US-10-750-185-47185

Query Match      82.2%  Score 14.8; DB 10; Length 3422;
Best Local Similarity 88.9%  Pred. No. 5.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1  CCCCCCTCTCTGGATAC 18
DB      2456  CCGCCCTCTCTGGATAC 2473

RESULT 50
US-10-750-623-47185
; Sequence 47185, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
```

```

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47185
; LENGTH: 3422
; TYPE: DNA
; ORGANISM: Bovine 19866881073042
US-10-750-623-47185

Query Match      82.2%  Score 14.8; DB 10; Length 3422;
Best Local Similarity 88.9%  Pred. No. 5.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1  CCCCCCTCTCTGGATAC 18
DB      2456  CCGCCCTCTCTGGATAC 2473

RESULT 51
US-10-330-773-911
; Sequence 911, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 911
; LENGTH: 69648
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-911

Query Match      82.2%  Score 14.8; DB 11; Length 69648;
Best Local Similarity 88.9%  Pred. No. 7.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1  CCCCCCTCTCTGGATAC 18
DB      21264  CCCACCTCTCTGGATCC 21281

RESULT 52
US-10-330-773-695
; Sequence 695, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 695
; LENGTH: 109147
; TYPE: DNA
; ORGANISM: Mus musculus
```

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)... (109147)  
OTHER INFORMATION: n = A,T,C or G  
US-10-330-773-695

Query Match      82.2%;    Score 14.8;    DB 11;    Length 109147;  
Best Local Similarity 88.9%;    Pred. No. 7.7e+02;  
Matches 16;    Conservative 0;    Mismatches 2;    Indels 0;    Gaps 0;

QY      1    CCCCCCTCTTCGTGATAC 18  
         |||||  
DB      11023    CCCCCGCTTCGTGATAC 11040

RESULT 53  
US-11-121-086-66  
; Sequence 66, Application US/11121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: POULSEN, TIM S.  
; APPLICANT: NIELSEN, KIRSTEN V.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-00000  
; CURRENT APPLICATION NUMBER: US/11/121,086  
; CURRENT FILING DATE: 2005-05-04  
; PRIOR APPLICATION NUMBER: 60/567,570  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 66  
; LENGTH: 163162  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-66

Query Match      82.2%;    Score 14.8;    DB 17;    Length 163162;  
Best Local Similarity 88.9%;    Pred. No. 8e+02;  
Matches 16;    Conservative 0;    Mismatches 2;    Indels 0;    Gaps 0;

QY      1    CCCCCCTCTTCGTGATAC 18  
         |||||  
DB      69041    CCCCCCTCTTCGTGATAC 69058

RESULT 54  
US-11-114-798-54/C  
; Sequence 54, Application US/11114798  
; Publication No. US20060035246A1  
; GENERAL INFORMATION:  
; APPLICANT: WU, KINA  
; APPLICANT: MARQUEZ, ABBEY  
; TITLE OF INVENTION: CHROMOGENIC IN SITU HYBRIDIZATION METHODS, KITS, AND  
; FILE REFERENCE: 0618.011.0004  
; CURRENT APPLICATION NUMBER: US/11/114,798  
; CURRENT FILING DATE: 2005-04-26  
; PRIOR APPLICATION NUMBER: 10/173,525  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 09/952,851  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn Ver. 3.3  
; SEQ ID NO 54  
; LENGTH: 198977  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-11-114-798-54

Query Match      82.2%;    Score 14.8;    DB 18;    Length 198977;

Best Local Similarity 88.9%;    Pred. No. 8.1e+02;  
Matches 16;    Conservative 0;    Mismatches 2;    Indels 0;    Gaps 0;

QY      1    CCCCCCTCTTCGTGATAC 18  
         |||||  
DB      127633    CCCCCCTCTTCGTGATAC 127616

RESULT 55  
US-09-824-575A-3/C  
; Sequence 3, Application US/09824575A  
; Publication No. US20060068382A1  
; GENERAL INFORMATION:  
; APPLICANT: BEASLEY, Ellen M. et al  
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
; FILE REFERENCE: CL000998  
; CURRENT APPLICATION NUMBER: US/09/824,575A  
; CURRENT FILING DATE: 2001-04-03  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 559686  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)... (559686)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-824-575A-3

Query Match      82.2%;    Score 14.8;    DB 8;    Length 559686;  
Best Local Similarity 88.9%;    Pred. No. 8.8e+02;  
Matches 16;    Conservative 0;    Mismatches 2;    Indels 0;    Gaps 0;

QY      1    CCCCCCTCTTCGTGATAC 18  
         |||||  
DB      49756    CCCCCCTCTTCGTGATAC 49739

RESULT 56  
US-10-995-561-49194  
; Sequence 49194, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 49194  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-995-561-49194

Query Match      80.0%;    Score 14.4;    DB 10;    Length 201;  
Best Local Similarity 93.8%;    Pred. No. 6.2e+02;  
Matches 15;    Conservative 0;    Mismatches 1;    Indels 0;    Gaps 0;

QY      1    CCCCCCTCTTCGTGAT 16  
         |||||  
DB      159    CCCCCCACTTCGTGAT 174

RESULT 57  
US-11-124-368A-13541  
; Sequence 13541, Application US/11124368A

```

; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13541
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-13541

```

```

Query Match      80.0%; Score 14.4; DB 17; Length 201;
Best Local Similarity 93.8%; Pred. No. 6.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 CCCCCCTCTCTGGAT 16
Db      159 CCCCCCACTTCTGGAT 174

```

```

RESULT 58
US-11-124-368A-13597
; Sequence 13597, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13597
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-13597

```

```

Query Match      80.0%; Score 14.4; DB 17; Length 201;
Best Local Similarity 93.8%; Pred. No. 6.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 CCCCCCTCTCTGGAT 16
Db      35 CCCCCCACTTCTGGAT 50

```

```

RESULT 59
US-11-124-367A-20109
; Sequence 20109, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with

```

```

; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20109
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-367A-20109

```

```

Query Match      80.0%; Score 14.4; DB 17; Length 201;
Best Local Similarity 93.8%; Pred. No. 6.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      3 CCCCCTCTCTGATAC 18
Db      13 CCACTCTTCTGATAC 28

```

```

RESULT 60
US-09-925-065A-478927/c
; Sequence 478927, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 478927
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-478927

```

```

Query Match      80.0%; Score 14.4; DB 7; Length 421;
Best Local Similarity 93.8%; Pred. No. 6.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      3 CCCCCTCTCTGATAC 18
Db      419 CCCCCTCTCTGACAC 404

```

```

RESULT 61
US-10-301-480-357016/c
; Sequence 357016, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome

```

```

; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 357016
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-357016

Query Match      80.0%; Score 14.4; DB 12; Length 479;
Best Local Similarity 93.8%; Pred. No. 6.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  CCCCTCTTCTGGATAC 18
      |  |||||
Db      309 CTCCTCTTCTGGATAC 294

RESULT 62
US-10-301-480-970425/c
; Sequence 970425, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 970425
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-970425

Query Match      80.0%; Score 14.4; DB 12; Length 479;
Best Local Similarity 93.8%; Pred. No. 6.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  CCCCTCTTCTGGATAC 18
      |  |||||
Db      309 CTCCTCTTCTGGATAC 294

RESULT 63
US-10-301-480-564341
; Sequence 564341, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; TYPE: DNA
```

```

; SEQ ID NO 564341
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-564341

Query Match      80.0%; Score 14.4; DB 12; Length 491;
Best Local Similarity 93.8%; Pred. No. 6.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  CCCCTCTTCTGGAT 16
      |  |||||
Db      384 CCCCTCTTCTGGAT 399

RESULT 64
US-10-301-480-1177750
; Sequence 1177750, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1177750
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1177750

Query Match      80.0%; Score 14.4; DB 12; Length 491;
Best Local Similarity 93.8%; Pred. No. 6.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  CCCCTCTTCTGGAT 16
      |  |||||
Db      384 CCCCTCTTCTGGAT 399

RESULT 65
US-09-925-065A-259422/c
; Sequence 259422, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 259422
; LENGTH: 523
; TYPE: DNA
```



```

; ORGANISM: Homo sapiens
; US-09-925-065A-259422

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 7; Length 523;
Pred. No. 7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCCTCTTGTGATAC 18
Db 117 CCCCTCTTGTGACAC 102

RESULT 66
US-09-925-065A-259423/c
; Sequence 259423, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 259423
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-259423

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 7; Length 523;
Pred. No. 7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCCTCTTGTGATAC 18
Db 117 CCCCTCTTGTGACAC 102

RESULT 67
US-09-925-065A-244698/c
; Sequence 244698, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 244698
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-244698

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 7; Length 529;
Pred. No. 7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCCTCTTGTGATAC 18
Db 117 CCCCTCTTGTGACAC 102

RESULT 68
US-10-301-480-337948/c
; Sequence 337948, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 337948
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-337948

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 12; Length 531;
Pred. No. 7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCCTCTTGTGATAC 18
Db 117 CCCCTCTTGTGACAC 102

RESULT 69
US-10-301-480-951357/c
; Sequence 951357, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 951357
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-951357

Query Match
Best Local Similarity 80.0%; Score 14.4; DB 12; Length 531;
Pred. No. 7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 3 CCCCTCTCTGGATAC 18  
 |||||  
 Db 117 CCCCTCTCTGGACAC 102

RESULT 70  
 US-09-925-065A-791113/c  
 ; Sequence 791113, Application US/09925065A  
 ; Publication No. US20040181048A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single  
 ; FILE REFERENCE: 108827.135  
 ; CURRENT APPLICATION NUMBER: US/09/925,065A  
 ; CURRENT FILING DATE: 2001-08-08  
 ; PRIOR APPLICATION NUMBER: US 60/243,096  
 ; PRIOR FILING DATE: 2000-10-24  
 ; PRIOR APPLICATION NUMBER: US 60/252,147  
 ; PRIOR FILING DATE: 2000-11-20  
 ; PRIOR APPLICATION NUMBER: US 60/250,092  
 ; PRIOR FILING DATE: 2000-11-30  
 ; PRIOR APPLICATION NUMBER: US 60/261,766  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/289,846  
 ; PRIOR FILING DATE: 2001-05-09  
 ; NUMBER OF SEQ ID NOS: 957086  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 791113  
 ; LENGTH: 545  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-925-065A-791113

Query Match 80.0%; Score 14.4; DB 7; Length 545;  
 Best Local Similarity 93.8%; Pred. No. 7e+02; 1; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 CCCCTCTCTGGATA 17  
 |||||  
 Db 73 CCCCTCTCTGTATA 58

Search completed: May 6, 2006, 19:28:41  
 Job time : 389 secs



PUBMED 11722936  
REFERENCE 2 (bases 1 to 1457)  
AUTHORS Burrell,P.C., Phalen,C.M. and Hovanec,T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria Group, 6100 Condon Dr., Moorpark, CA 93021, USA

FEATURES  
source  
1.1457  
/organism="Nitrosomonas sp. R7c140"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated from freshwater aquaria enrichments"  
/db\_xref="taxon:160543"  
/clone="R7c140"  
/environmental\_sample  
<1..>1457  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 100.0%; Score 1457; DB 3; Length 1457;  
Beet Local Similarity 100.0%; Pred. No. 8.3e-16;  
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGCTGCGGCGATCTTTACATGCAAGTCGAACGGCAGACGGATGCTTGCA 60  
DB 1 ATTGAACGCTGCGGCGATCTTTACATGCAAGTCGAACGGCAGACGGATGCTTGCA 60  
QY 61 CTGTGGCGAGTGGCGGAGCGGCTGAGTAATGCAATCGGAACGATTCAGAAAGGGGGGTA 120  
DB 61 CTGTGGCGAGTGGCGGAGCGGCTGAGTAATGCAATCGGAACGATTCAGAAAGGGGGGTA 120  
QY 121 ACGCATCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCGGGGATCGAAA 180  
DB 121 ACGCATCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCGGGGATCGAAA 180  
QY 181 GACCTTGGCGCTTTTGAAGCGCGCATGTCTGATTAGCTGATGGGTTAAAGGCTTAC 240  
DB 181 GACCTTGGCGCTTTTGAAGCGCGCATGTCTGATTAGCTGATGGGTTAAAGGCTTAC 240  
QY 241 CAAGCGAGCATCAGTATGTTGGTCTTGAGAGAGCAGACAGCCACACTGGGACTGAGACAG 300  
DB 241 CAAGCGAGCATCAGTATGTTGGTCTTGAGAGAGCAGACAGCCACACTGGGACTGAGACAG 300  
QY 301 GCCCAGACTCTTACCGGAGGCGAGCATGTGGGGAATTTTGAACAATGGGCGCAAGCTGATC 360  
DB 301 GCCCAGACTCTTACCGGAGGCGAGCATGTGGGGAATTTTGAACAATGGGCGCAAGCTGATC 360  
QY 361 CAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTTTAAAGCTCTTTCAATCGAGAGAGA 420  
DB 361 CAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTTTAAAGCTCTTTCAATCGAGAGAGA 420  
QY 421 AAAGGTTACGATTAATATGTAATCATGACGCGTATCGACAGAGAGACACCGGCTAAC 480  
DB 421 AAAGGTTACGATTAATATGTAATCATGACGCGTATCGACAGAGAGACACCGGCTAAC 480  
QY 481 TACGTGCGCAGACGCGCGGTATACGTAGGGTGCAAGCGTTTATCGAATTTACTGGGCGT 540  
DB 481 TACGTGCGCAGACGCGCGGTATACGTAGGGTGCAAGCGTTTATCGAATTTACTGGGCGT 540  
QY 541 AAAGGTTGCGAGCGGCTTTTGAAGTCAATGTAATCCCGGCTTAACTGGGAGAT 600  
DB 541 AAAGGTTGCGAGCGGCTTTTGAAGTCAATGTAATCCCGGCTTAACTGGGAGAT 600  
QY 601 TCGGTTTGAACATCAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTATGACAG 660  
DB 601 TCGGTTTGAACATCAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTATGACAG 660  
QY 661 TGAATTCGCTAGATATGAAAGACATGATGAGGCGAAGCGCTCCCTGGGTTAACACT 720  
DB 661 TGAATTCGCTAGATATGAAAGACATGATGAGGCGAAGCGCTCCCTGGGTTAACACT 720  
QY 721 GACGCTCATGACGAAGCGTGGGAGCAACAGATTAGATACCTGTAAGTCCAGCGC 780  
DB 721 GACGCTCATGACGAAGCGTGGGAGCAACAGATTAGATACCTGTAAGTCCAGCGC 780

DB 721 GACGCTCATGACGAAGCGTGGGAGCAACAGATTAGATACCTGTAAGTCCAGCGC 780  
QY 781 CTTAAACGATGCACTAATGTTGGGCTTTATGAGCTTGTGTAACGAAGCTAACGGCGGA 840  
DB 781 CTTAAACGATGCACTAATGTTGGGCTTTATGAGCTTGTGTAACGAAGCTAACGGCGGA 840  
QY 841 AGTTGACCGCTGGGAGATACGCTGCGCAAGATTAATACTCAAGAGATTTGACGGGAGCC 900  
DB 841 AGTTGACCGCTGGGAGATACGCTGCGCAAGATTAATACTCAAGAGATTTGACGGGAGCC 900  
QY 901 GCACAGGCGGTGATATGATGATTAATTCATGCAACGCGAATACTTACCTTACCTT 960  
DB 901 GCACAGGCGGTGATATGATGATTAATTCATGCAACGCGAATACTTACCTTACCTT 960  
QY 961 GACATGTAGCGAATTTTGAAGATAGATTAGTGTGCGGAAAGCTTAAACAGGTCGTG 1020  
DB 961 GACATGTAGCGAATTTTGAAGATAGATTAGTGTGCGGAAAGCTTAAACAGGTCGTG 1020  
QY 1021 CATGCTGTGTCAGCTCGTGTGAGATGTTGGGTTAAGTCCGCAACGACGCAAC 1080  
DB 1021 CATGCTGTGTCAGCTCGTGTGAGATGTTGGGTTAAGTCCGCAACGACGCAAC 1080  
QY 1081 CTGTGATTAATTCGATCAATTTGGTGGCACTTTAATGAGACTGCCGTGACAAACCG 1140  
DB 1081 CTGTGATTAATTCGATCAATTTGGTGGCACTTTAATGAGACTGCCGTGACAAACCG 1140  
QY 1141 GAGAAAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
DB 1141 GAGAAAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
QY 1201 TACAATGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
DB 1201 TACAATGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
QY 1261 TCGTATGTCGAGATGGAATGCTGCAATCTGACTCCGTGAATGCGAATGCTAATATGCC 1320  
DB 1261 TCGTATGTCGAGATGGAATGCTGCAATCTGACTCCGTGAATGCGAATGCTAATATGCC 1320  
QY 1321 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
DB 1321 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
QY 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTTAAACGTAAGAGGCGCTTCCACAGCTGA 1440  
DB 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTTAAACGTAAGAGGCGCTTCCACAGCTGA 1440  
QY 1441 GATTGATGATGAGGCTG 1457  
DB 1441 GATTGATGATGAGGCTG 1457

RESULT 2  
LOCUS CQ796908 1457 bp DNA linear PAT 19-APR-2004  
DEFINITION Sequence 1 from Patent WO2004026772.  
ACCESSION CQ796908  
VERSION CQ796908.1 GI:46408534  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1  
AUTHORS Hovanec,T.A.  
TITLE Ammonia-oxidizing bacteria and methods of using and detecting  
JOURNAL thesame  
Patent: WO 2004026772-A 1 01-APR-2004;  
Aquadria Inc. (US)  
FEATURES  
source  
1.1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="AOB Type A R7c1ome140 16S rDNA"

ORIGIN

Query Match 100.0%; Score 1457; DB 6; Length 1457;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-16;  
 Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATTGAACGCTGGCGGCATCTTTTACATGCAAGTCGAAACGCGACAGACGGATGCTTGCAT 60  
 1 ATTGAACGCTGGCGGCATCTTTTACATGCAAGTCGAAACGCGACAGACGGATGCTTGCAT 60  
 61 CTGGTGGGAGATGGGGGACGGGTGATATGATGGGACGATATCCAGAAAGAGGGGGGTA 120  
 61 CTGGTGGGAGATGGGGGACGGGTGATATGATGGGACGATATCCAGAAAGAGGGGGGTA 120  
 61 CTGGTGGGAGATGGGGGACGGGTGATATGATGGGACGATATCCAGAAAGAGGGGGGTA 120  
 121 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAAGAGAGAAAGAGGGGATCGAAA 180  
 121 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAAGAGAGAAAGAGGGGATCGAAA 180  
 181 GACCTTGGCGCTTTTGGAGCGCGCGATGTCTGATTAGCTAGTTGTGGGTAAAGGCTTAC 240  
 181 GACCTTGGCGCTTTTGGAGCGCGCGATGTCTGATTAGCTAGTTGTGGGTAAAGGCTTAC 240  
 241 CAAGCGAGATGATGATGTTGCTGAGAGAGACAGCAGCCACCTGGGACTGAGACAG 300  
 241 CAAGCGAGATGATGATGTTGCTGAGAGAGACAGCAGCCACCTGGGACTGAGACAG 300  
 241 CAAGCGAGATGATGATGTTGCTGAGAGAGACAGCAGCCACCTGGGACTGAGACAG 300  
 301 GCCCAAGCTCTACCGGAGGACAGAGTGGGAAATTTTGGACAATGGGCGCAAGCTGATC 360  
 301 GCCCAAGCTCTACCGGAGGACAGAGTGGGAAATTTTGGACAATGGGCGCAAGCTGATC 360  
 301 GCCCAAGCTCTACCGGAGGACAGAGTGGGAAATTTTGGACAATGGGCGCAAGCTGATC 360  
 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCGGGTGTAAGCTCTTTCAGTCGAGAGA 420  
 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCGGGTGTAAGCTCTTTCAGTCGAGAGA 420  
 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCGGGTGTAAGCTCTTTCAGTCGAGAGA 420  
 421 AAAAGTTACGTTAATATCTGTAATCTGATGACGCTATCCAGAGAAAGACCGGCTAAC 480  
 421 AAAAGTTACGTTAATATCTGTAATCTGATGACGCTATCCAGAGAAAGACCGGCTAAC 480  
 421 AAAAGTTACGTTAATATCTGTAATCTGATGACGCTATCCAGAGAAAGACCGGCTAAC 480  
 481 TACGTGCCAGACGCGCGGTATACGTAAGGTCGAAAGCTTAACTGAGGCGT 540  
 481 TACGTGCCAGACGCGCGGTATACGTAAGGTCGAAAGCTTAACTGAGGCGT 540  
 481 TACGTGCCAGACGCGCGGTATACGTAAGGTCGAAAGCTTAACTGAGGCGT 540  
 541 AAAAGGTGCGCAGCGCGCTTTTGAAGTCAGATGTAATCCCGGCGCTTAACTGCGAAT 600  
 541 AAAAGGTGCGCAGCGCGCTTTTGAAGTCAGATGTAATCCCGGCGCTTAACTGCGAAT 600  
 541 AAAAGGTGCGCAGCGCGCTTTTGAAGTCAGATGTAATCCCGGCGCTTAACTGCGAAT 600  
 601 TGGCTTTGAAATCAAGGCTTAAAGTGTGCAAGAGGAGTGGAAATTCATGTGTAGCAG 660  
 601 TGGCTTTGAAATCAAGGCTTAAAGTGTGCAAGAGGAGTGGAAATTCATGTGTAGCAG 660  
 601 TGGCTTTGAAATCAAGGCTTAAAGTGTGCAAGAGGAGTGGAAATTCATGTGTAGCAG 660  
 661 TGAATATGCGATGATATGGAAGAAATCATGATGGCGAAGGACGCTCTGGTTAACT 720  
 661 TGAATATGCGATGATATGGAAGAAATCATGATGGCGAAGGACGCTCTGGTTAACT 720  
 661 TGAATATGCGATGATATGGAAGAAATCATGATGGCGAAGGACGCTCTGGTTAACT 720  
 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAATACCTGTGATGTCAGGCC 780  
 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAATACCTGTGATGTCAGGCC 780  
 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAATACCTGTGATGTCAGGCC 780  
 781 CTAAACGATGTCAACTAGTTGTTGGGCTTATTAAGGCTTGTGTAACGAAAGCTAACGCTGA 840  
 781 CTAAACGATGTCAACTAGTTGTTGGGCTTATTAAGGCTTGTGTAACGAAAGCTAACGCTGA 840  
 781 CTAAACGATGTCAACTAGTTGTTGGGCTTATTAAGGCTTGTGTAACGAAAGCTAACGCTGA 840  
 841 AGTTGACCGCTGGGGGAGTACGGTCCGAAAGTTAAACTCAAGAGATTTAGCGGGAGCC 900  
 841 AGTTGACCGCTGGGGGAGTACGGTCCGAAAGTTAAACTCAAGAGATTTAGCGGGAGCC 900  
 841 AGTTGACCGCTGGGGGAGTACGGTCCGAAAGTTAAACTCAAGAGATTTAGCGGGAGCC 900  
 901 GCAACAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAACCTTACCTTACCTT 960  
 901 GCAACAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAACCTTACCTTACCTT 960  
 901 GCAACAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAACCTTACCTTACCTT 960  
 961 GACATGTAGGAATTTTCTAGAGATTAATAGTCTTCGGGAAAGCTTAACAGAGTCTG 1020  
 961 GACATGTAGGAATTTTCTAGAGATTAATAGTCTTCGGGAAAGCTTAACAGAGTCTG 1020  
 961 GACATGTAGGAATTTTCTAGAGATTAATAGTCTTCGGGAAAGCTTAACAGAGTCTG 1020

1021 CATGCTGTGCTGACGCTTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1080  
 1021 CATGCTGTGCTGACGCTTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1080  
 1081 CTGTGCTTAATATGTCATCATTTTGGTGGGACCTTAAATGAGATGTCGGGTGACAAACCG 1140  
 1081 CTGTGCTTAATATGTCATCATTTTGGTGGGACCTTAAATGAGATGTCGGGTGACAAACCG 1140  
 1081 CTGTGCTTAATATGTCATCATTTTGGTGGGACCTTAAATGAGATGTCGGGTGACAAACCG 1140  
 1141 GAGGAGGTGGGAGATGACGTCAAGTCTTCATGCGCTTATGGGTAGGGCTTACACGTAA 1200  
 1141 GAGGAGGTGGGAGATGACGTCAAGTCTTCATGCGCTTATGGGTAGGGCTTACACGTAA 1200  
 1141 GAGGAGGTGGGAGATGACGTCAAGTCTTCATGCGCTTATGGGTAGGGCTTACACGTAA 1200  
 1201 TACAAATGCGGCTACAGAGGGTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
 1201 TACAAATGCGGCTACAGAGGGTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
 1201 TACAAATGCGGCTACAGAGGGTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
 1261 TCGTAGTCCGATCGGAATCTGCAACTGCACTCCGTGAAGTGGAAATGCTAATATGCG 1320  
 1261 TCGTAGTCCGATCGGAATCTGCAACTGCACTCCGTGAAGTGGAAATGCTAATATGCG 1320  
 1261 TCGTAGTCCGATCGGAATCTGCAACTGCACTCCGTGAAGTGGAAATGCTAATATGCG 1320  
 1321 GGATCAGCATGTCCGCTGATATAGTTCCTCGGCTTTGTAACACCCGCTTACACAT 1380  
 1321 GGATCAGCATGTCCGCTGATATAGTTCCTCGGCTTTGTAACACCCGCTTACACAT 1380  
 1321 GGATCAGCATGTCCGCTGATATAGTTCCTCGGCTTTGTAACACCCGCTTACACAT 1380  
 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTTAAACGTAAGAGAGGCGCTTGGCAGGTGA 1440  
 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTTAAACGTAAGAGAGGCGCTTGGCAGGTGA 1440  
 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTTAAACGTAAGAGAGGCGCTTGGCAGGTGA 1440  
 1441 GATTATGATCTGGGGTG 1457  
 1441 GATTATGATCTGGGGTG 1457  
 1441 GATTATGATCTGGGGTG 1457

RESULT 3  
 CS089154 1457 bp DNA linear PAT 25-MAY-2005  
 LOCUS CS089154 Sequence 1 from Patent EP1502948.  
 DEFINITION CS089154  
 ACCESSION CS089154  
 VERSION CS089154.1 GI:66714438  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 unclassified  
 unclassified  
 unclassified  
 unclassified

REFERENCE  
 1 Hovanec,T.A. and Burrell,P.C.  
 Ammonia-oxidizing bacteria  
 Patent: EP 1502948-A 1 02-FEB-2005;  
 Aquaria Inc. (US)

FEATURES  
 source  
 1. 1457  
 /organism="unclassified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"  
 /note="Description of unknown Organism: AOB Type A  
 ammonia-oxidizing bacterium represented by R7 clone140"

ORIGIN

Query Match 100.0%; Score 1457; DB 6; Length 1457;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-16;  
 Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATTGAACGCTGGCGGCATCTTTTACATGCAAGTCGAAACGCGACAGACGGATGCTTGCAT 60  
 1 ATTGAACGCTGGCGGCATCTTTTACATGCAAGTCGAAACGCGACAGACGGATGCTTGCAT 60  
 61 CTGGTGGGAGATGGGGGACGGGTGATATGATGGGACGATATCCAGAAAGAGGGGGGTA 120  
 61 CTGGTGGGAGATGGGGGACGGGTGATATGATGGGACGATATCCAGAAAGAGGGGGGTA 120  
 61 CTGGTGGGAGATGGGGGACGGGTGATATGATGGGACGATATCCAGAAAGAGGGGGGTA 120  
 1201 TACAAATGCGGCTACAGAGGGTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
 1201 TACAAATGCGGCTACAGAGGGTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
 1201 TACAAATGCGGCTACAGAGGGTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
 1261 TCGTAGTCCGATCGGAATCTGCAACTGCACTCCGTGAAGTGGAAATGCTAATATGCG 1320  
 1261 TCGTAGTCCGATCGGAATCTGCAACTGCACTCCGTGAAGTGGAAATGCTAATATGCG 1320  
 1261 TCGTAGTCCGATCGGAATCTGCAACTGCACTCCGTGAAGTGGAAATGCTAATATGCG 1320  
 1321 GGATCAGCATGTCCGCTGATATAGTTCCTCGGCTTTGTAACACCCGCTTACACAT 1380  
 1321 GGATCAGCATGTCCGCTGATATAGTTCCTCGGCTTTGTAACACCCGCTTACACAT 1380  
 1321 GGATCAGCATGTCCGCTGATATAGTTCCTCGGCTTTGTAACACCCGCTTACACAT 1380  
 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTTAAACGTAAGAGAGGCGCTTGGCAGGTGA 1440  
 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTTAAACGTAAGAGAGGCGCTTGGCAGGTGA 1440  
 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTTAAACGTAAGAGAGGCGCTTGGCAGGTGA 1440  
 1441 GATTATGATCTGGGGTG 1457  
 1441 GATTATGATCTGGGGTG 1457  
 1441 GATTATGATCTGGGGTG 1457

QY 181 GACCTTGGCTTTTGGAGCGCCGAGTGTCTAGTTAGTTGTTGGGTTAAAGGCTTAC 240  
DB 181 GACCTTGGCGCTTTTGGAGCGCCGAGTGTCTAGTTAGTTGTTGGGTTAAAGGCTTAC 240  
QY 241 CAAGCGACGATCAGTAGTTGGTCTGAGAGACGACACGACCTAGTGGACTGAGACAG 300  
DB 241 CAAGCGACGATCAGTAGTTGGTCTGAGAGACGACACGACCTAGTGGACTGAGACAG 300  
QY 301 GCCGAGCTCTTACGGGAGGCGACAGTGGGGAATTTTGGACAAATGCGCGCAAGCTGATC 360  
DB 301 GCCGAGCTCTTACGGGAGGCGACAGTGGGGAATTTTGGACAAATGCGCGCAAGCTGATC 360  
QY 361 CAGCAATGCCGCTGAGTGAAGAAAGGCTTGGGTTTGAATCTTTTCAGTGGAGAA 420  
DB 361 CAGCAATGCCGCTGAGTGAAGAAAGGCTTGGGTTTGAATCTTTTCAGTGGAGAA 420  
QY 421 AAAGGTTACGTAATTAATCTGACCTGACATGACGATGCAAGAGAAAGACCGGCTTAC 480  
DB 421 AAAGGTTACGTAATTAATCTGACCTGACATGACGATGCAAGAGAAAGACCGGCTTAC 480  
QY 481 TACGTCGACAGCGCGGCTTAACTGAGGGGTGCAAGCGTTAATCGGAATTTCTGGCGT 540  
DB 481 TACGTCGACAGCGCGGCTTAACTGAGGGGTGCAAGCGTTAATCGGAATTTCTGGCGT 540  
QY 541 AAAGGTTGCGGAGGCGGCTTGTAGTCAATGATGAAATCCCGGGCTTAACTGGGAT 600  
DB 541 AAAGGTTGCGGAGGCGGCTTGTAGTCAATGATGAAATCCCGGGCTTAACTGGGAT 600  
QY 601 TCGCTTTGAACTCAAGAGCTAGAGTGTGGCAGAGAGGTGGAATTCATGTAGCAG 660  
DB 601 TCGCTTTGAACTCAAGAGCTAGAGTGTGGCAGAGAGGTGGAATTCATGTAGCAG 660  
QY 661 TGAATTCGCTAGAGATATGGAAGAAATCTGATGCGGAGGCGACCTCTGGGTTAACT 720  
DB 661 TGAATTCGCTAGAGATATGGAAGAAATCTGATGCGGAGGCGACCTCTGGGTTAACT 720  
QY 721 GACGCTCATGCGAAGAGGTGGGAGGAGCAAGATTAATACCTGGTGTAGTCCAGCGC 780  
DB 721 GACGCTCATGCGAAGAGGTGGGAGGAGCAAGATTAATACCTGGTGTAGTCCAGCGC 780  
QY 781 CTAAACGATGTCAATAGTTGTGTGGCCCTTATTAGGCTTGTGAAGAACTAACCGCTGA 840  
DB 781 CTAAACGATGTCAATAGTTGTGTGGCCCTTATTAGGCTTGTGAAGAACTAACCGCTGA 840  
QY 841 AGTTGACCGCTGGGAGAGTACGCTGCGCAAGATTTAAACTCAAGAAATTGACGGGAGCC 900  
DB 841 AGTTGACCGCTGGGAGAGTACGCTGCGCAAGATTTAAACTCAAGAAATTGACGGGAGCC 900  
QY 901 GCACAAAGGCGTGGATTAATGTGATTAATTCGATGCAACCGGAAACCTTACCTTACCCTT 960  
DB 901 GCACAAAGGCGTGGATTAATGTGATTAATTCGATGCAACCGGAAACCTTACCTTACCCTT 960  
QY 961 GACATGTAGCGAATTTTCTAGAGATAGATAGTGTGGGAAACGCTAACAGGTGCTG 1020  
DB 961 GACATGTAGCGAATTTTCTAGAGATAGATAGTGTGGGAAACGCTAACAGGTGCTG 1020  
QY 1021 CATGCTGTGCTGACGCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCCCAAC 1080  
DB 1021 CATGCTGTGCTGACGCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCCCAAC 1080  
QY 1081 CTGTGCTAATTAATGATCAATTTGGTGGGCACTTAAATGAGACTGCGGGTGAACAAACG 1140  
DB 1081 CTGTGCTAATTAATGATCAATTTGGTGGGCACTTAAATGAGACTGCGGGTGAACAAACG 1140  
QY 1141 GAGGAAGGTGGGATGAGCTCAAGTCTCATGCGCCCTTATGGGTAGGGCTTTCACAGTGA 1200  
DB 1141 GAGGAAGGTGGGATGAGCTCAAGTCTCATGCGCCCTTATGGGTAGGGCTTTCACAGTGA 1200  
QY 1201 TACAAATGCGGCTACAGAGGTTGCGCAACCGCGAGGGGAGACTTATCTCAGAAAGCGCG 1260  
DB 1201 TACAAATGCGGCTACAGAGGTTGCGCAACCGCGAGGGGAGACTTATCTCAGAAAGCGCG 1260

QY 1261 TCGTAGTCCGAGTCGAGTCTGCAACTGCACTCCGTGAAAGTCGGAATCGTAGTAATCGC 1320  
DB 1261 TCGTAGTCCGAGTCGAGTCTGCAACTGCACTCCGTGAAAGTCGGAATCGTAGTAATCGC 1320  
QY 1321 GGATCAGCATGTCCGGGTAATACGTTCCGGGCTTGTACACACCGCCGCTCACACCAT 1380  
DB 1321 GGATCAGCATGTCCGGGTAATACGTTCCGGGCTTGTACACACCGCCGCTCACACCAT 1380  
QY 1381 GGGAGTGGGTTTTCACAGAAACAGTAGTCTTAACCTTAAGAGAGGCGCTTCCACAGTGA 1440  
DB 1381 GGGAGTGGGTTTTCACAGAAACAGTAGTCTTAACCTTAAGAGAGGCGCTTCCACAGTGA 1440  
QY 1441 GATTTCATGACTGGGGTG 1457  
DB 1441 GATTTCATGACTGGGGTG 1457

RESULT 4  
CS089155  
LOCUS  
DEFINITION Sequence 2 from Patent EP1502948.  
ACCESSION CS089155  
VERSION CS089155.1 GI:66714439  
KEYWORDS  
SOURCE  
ORGANISM  
unidentified  
unclassified.

REFERENCE  
1 Hovaneec, T. A. and Burrell, P. C.  
AUTHORS  
TITLE  
JOURNAL  
Patent: EP 1502948-A 2 02-FEB-2005;  
Aquaia Inc. (US)

FEATURES  
source  
location/qualifiers  
1..1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="Description of Unknown Organism: AOB Type A1  
amonnia-oxidizing bacterium represented by R7 clone187"

ORIGIN

Query Match 100.0%; Score 1457; DB 6; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 8.3e-16;  
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTGAACGCTGGCGGATGCTTTTACATGCAAGTCCGACGACGAGATGCTTGCAT 60  
DB 1 ATTGAACGCTGGCGGATGCTTTTACATGCAAGTCCGACGACGAGATGCTTGCAT 60  
QY 61 CTGGTGGCGAGTGGCGGAGCGGGTGAATGATGATCGGAACGATCCAGAAAGGCGGTA 120  
DB 61 CTGGTGGCGAGTGGCGGAGCGGGTGAATGATGATCGGAACGATCCAGAAAGGCGGTA 120  
QY 121 ACGCATGGAAGATGTGTCTAATACCGCATATACTTAAAGAGAAAGAGGAGATCGAAA 180  
DB 121 ACGCATGGAAGATGTGTCTAATACCGCATATACTTAAAGAGAAAGAGGAGATCGAAA 180  
QY 181 GACCTTGGGCTTTTGGAGCGGCGCATGTCTAATTAAGTGAAGTGGTGGGTTAAAGCCCTAC 240  
DB 181 GACCTTGGGCTTTTGGAGCGGCGCATGTCTAATTAAGTGAAGTGGTGGGTTAAAGCCCTAC 240  
QY 241 CAAGCGACGATCAGTAGTTGGTCTGAGAGGACGACACGACCTGGGACTGAGACAG 300  
DB 241 CAAGCGACGATCAGTAGTTGGTCTGAGAGGACGACACGACCTGGGACTGAGACAG 300  
QY 301 GCCGAGACTCTTACGGGAGGCGAGCAGTGGGGAATTTTGGACAAATGCGCGCAAGCTGATC 360  
DB 301 GCCGAGACTCTTACGGGAGGCGAGCAGTGGGGAATTTTGGACAAATGCGCGCAAGCTGATC 360  
QY 361 CAGCAATGCCGCTGAGTGAAGAAAGGCTTGGGTTTGAATCTTTTCAGTGGAGAA 420  
DB 361 CAGCAATGCCGCTGAGTGAAGAAAGGCTTGGGTTTGAATCTTTTCAGTGGAGAA 420

QY	421	AAAGGTTACGGTAATTAATCGTGA	CTCATGACGGTATTCGACAGAAAGAAAGACCGGGCTAAC	480
Db	421	AAAGGTTACGGTAATTAATCGTGA	CTCATGACGGTATTCGACAGAAAGAAAGACCGGGCTAAC	480
QY	481	TACGTGCAGACGCGCGGTAAATCGT	AGGGGTGCAAGCGTTAATCGAATTTACTGGGCGT	540
Db	481	TACGTGCAGACGCGCGGTAAATCGT	AGGGGTGCAAGCGTTAATCGAATTTACTGGGCGT	540
QY	541	AAAGGTCGCGACGCGCTTTGTAA	AGTCAGATGTGAATCCCGGGCTTAACTTGGGAAT	600
Db	541	AAAGGTCGCGACGCGCTTTGTAA	AGTCAGATGTGAATCCCGGGCTTAACTTGGGAAT	600
QY	601	TGCGTTTGAATCTACAAAGGCTA	AGAGTGTGGCAGAGGAGGTGTGAATTCATGTGTACAG	660
Db	601	TGCGTTTGAATCTACAAAGGCTA	AGAGTGTGGCAGAGGAGGTGTGAATTCATGTGTACAG	660
QY	661	TGAAATGCGTGAAGATATGGAAGAA	CATCGATGGCGAAGGACGCTCCGGGTTAACTACT	720
Db	661	TGAAATGCGTGAAGATATGGAAGAA	CATCGATGGCGAAGGACGCTCCGGGTTAACTACT	720
QY	721	GACGCTCATGACAGAAACGCTGGG	AGGACAAACAGATTAGTACCTGTAGTACAGCC	780
Db	721	GACGCTCATGACAGAAACGCTGGG	AGGACAAACAGATTAGTACCTGTAGTACAGCC	780
QY	781	CTAAACGATGTCAACTAGTGTGGG	CCCTTATAGCTTGGTAAAGAACTTAACGCTGA	840
Db	781	CTAAACGATGTCAACTAGTGTGGG	CCCTTATAGCTTGGTAAAGAACTTAACGCTGA	840
QY	841	AGTTGACCGCTCGGGGAATACGGT	CGGAAGTTAAACTCAABGAATGACGGGGACCC	900
Db	841	AGTTGACCGCTCGGGGAATACGGT	CGGAAGTTAAACTCAABGAATGACGGGGACCC	900
QY	901	GCACAAGCGGTGAGATTATGTGGA	TTAATTCATGCAACGCCAAAAACCTTACTTACCCTT	960
Db	901	GCACAAGCGGTGAGATTATGTGGA	TTAATTCATGCAACGCCAAAAACCTTACTTACCCTT	960
QY	961	GACATGTAGCGAATTTCTAAGATAG	ATTAGTGTCTGGGAAGGCTTAACACAGGTGCTG	1020
Db	961	GACATGTAGCGAATTTCTAAGATAG	ATTAGTGTCTGGGAAGGCTTAACACAGGTGCTG	1020
QY	1021	CATGGCTGTCTGACGCTGTGTCTG	CGGAANTGTGGGTTAAGTCCCGGACAGAGGGCAACC	1080
Db	1021	CATGGCTGTCTGACGCTGTGTCTG	CGGAANTGTGGGTTAAGTCCCGGACAGAGGGCAACC	1080
QY	1081	CTTGTCAATTAATTTGCATCATTT	TGGTGGGCACTTAAATGAGACTGCGGTGACAAACCG	1140
Db	1081	CTTGTCAATTAATTTGCATCATTT	TGGTGGGCACTTAAATGAGACTGCGGTGACAAACCG	1140
QY	1141	GAGGAAGGTGGGAGTAGCGTCAAG	TCTCATAGGCCCTTATATGGGTATGGGCTTCAACGTAA	1200
Db	1141	GAGGAAGGTGGGAGTAGCGTCAAG	TCTCATAGGCCCTTATATGGGTATGGGCTTCAACGTAA	1200
QY	1201	TACAAATGGCGGCTACAGAGGTT	GCCAACCCGCGAGGGGAGCTTAATCTCAGAAAGCCGC	1260
Db	1201	TACAAATGGCGGCTACAGAGGTT	GCCAACCCGCGAGGGGAGCTTAATCTCAGAAAGCCGC	1260
QY	1261	TGCTATGTCGGATTCGGAAGTCTG	CAACTGGAATCCGATGAGTGTGGAATTCGCTATGATATGC	1320
Db	1261	TGCTATGTCGGATTCGGAAGTCTG	CAACTGGAATCCGATGAGTGTGGAATTCGCTATGATATGC	1320
QY	1321	GGATCAGCATGTCTGGCGTGAATAC	GTTCGCGGCTTGTATACACACGCCCCGTCAACCAT	1380
Db	1321	GGATCAGCATGTCTGGCGTGAATAC	GTTCGCGGCTTGTATACACACGCCCCGTCAACCAT	1380
QY	1381	GGGAGTGGGTTTCAACAGAGCAG	GTAGTCTTAACCGTAAAGAGGGGCGCTTGCCACGCTGA	1440
Db	1381	GGGAGTGGGTTTCAACAGAGCAG	GTAGTCTTAACCGTAAAGAGGGGCGCTTGCCACGCTGA	1440
QY	1441	GATTCAATGATGGGGTG	1457	
Db	1441	GATTCAATGATGGGGTG	1457	

AX316092	AX316092	1457 bp	DNA	linear	PAT 14-DEC-2001
LOCUS	AX316092	Sequence 1 from Patent WO0190312.			
DEFINITION	AX316092				
ACCESSION	AX316092.1	GI:17899283			
VERSION					
KEYWORDS					
SOURCE	unidentified				
ORGANISM	unclassified sequences.				
REFERENCE	1				
AUTHORS	Hovanec, T.A. and Burrell, P.C.				
TITLE	Ammonia-oxidizing bacteria				
JOURNAL	Patent: WO 0190312-A 1 29-NOV-2001;				
	AQUARIA, INC. (US)				
FEATURES	Location/Qualifiers				
source	1..1457				
	/organism="unidentified"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32644"				
	/note="AOB Type A ammonia-oxidizing bacterium represented by R7 clone140"				
ORIGIN					
Query Match	100.0%;	Score 1457;	DB 6;	Length 1457;	
Best Local Similarity	100.0%;	Pred. No. 8.3e-16;			
Matches 1457;	Conservative	0;	Mismatches	0;	Indels
					Gaps
					0;
QY	1	ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAAACGGCAGCAGATGCTTGAT	60		
DB	1	ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAAACGGCAGCAGATGCTTGAT	60		
QY	61	CTGTGGCGAGTGGCGGACGGGTGAGTATGATGTAAGCTATTCAGAAAGGGGGGTA	120		
DB	61	CTGTGGCGAGTGGCGGACGGGTGAGTATGATGTAAGCTATTCAGAAAGGGGGGTA	120		
QY	121	ACGCATCGAAAGATGCTTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAA	180		
DB	121	ACGCATCGAAAGATGCTTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAA	180		
QY	181	GACCTTGGGCTTTGGAGCGGCGCCAGTGTCTGATTAGTCTAGTTGGTGGGGTAAAGGCTTAC	240		
DB	181	GACCTTGGGCTTTGGAGCGGCGCCAGTGTCTGATTAGTCTAGTTGGTGGGGTAAAGGCTTAC	240		
QY	241	CAGGCGACGATCAGTAGTGGTCTGAGAGACGACACCACTGGGACTGAGACACG	300		
DB	241	CAGGCGACGATCAGTAGTGGTCTGAGAGACGACACCACTGGGACTGAGACACG	300		
QY	301	GCCCGACCTCTTAACGGGAGGCGACGAGTGGGGAATTTGACAAATGGGGCGAAGCCTGATC	360		
DB	301	GCCCGACCTCTTAACGGGAGGCGACGAGTGGGGAATTTGACAAATGGGGCGAAGCCTGATC	360		
QY	361	CAGCAATGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGGCTTTTCAGTCGAGAAGA	420		
DB	361	CAGCAATGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGGCTTTTCAGTCGAGAAGA	420		
QY	421	AAAGGTTACGTTAAATTAATCTGATCTGACGATGACAGAAAGAACACCGGCTAAC	480		
DB	421	AAAGGTTACGTTAAATTAATCTGATCTGACGATGACAGAAAGAACACCGGCTAAC	480		
QY	481	TACGTGCGACGACCGCGGTAATACGTAGGGTGCAGAGCTTAATCGGAATTAATCTGGGCT	540		
DB	481	TACGTGCGACGACCGCGGTAATACGTAGGGTGCAGAGCTTAATCGGAATTAATCTGGGCT	540		
QY	541	AAAGGTGGCGAGCGGCTTTGTAAGTCAAGATGAATCCCGGGCTTAACCTGGGAAT	600		
DB	541	AAAGGTGGCGAGCGGCTTTGTAAGTCAAGATGAATCCCGGGCTTAACCTGGGAAT	600		
QY	601	TGGGTTTGAACCTACAGGCTAGATGTGGCAGAGGAGTGGAAATTCATGTGTAGCAG	660		
DB	601	TGGGTTTGAACCTACAGGCTAGATGTGGCAGAGGAGTGGAAATTCATGTGTAGCAG	660		
QY	661	TGAAATGCGTAGAATATGAAGAACATCATGTCGGAAGCAGCCTCTCGGGTTAACT	720		

Db 661 TGAATTCGGTGAAGATATGAAAGAAATGATGCGGAAGCGACCTCTCGGTTAACT 720  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTTGGTAGTCCAGCC 780  
Db 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTTGGTAGTCCAGCC 780  
Qy 781 CTAAACGATGCTAACTAGTGTGGGCTTATTAGCTTGTGAAGAACTAAACCGCTGA 840  
Db 781 CTAAACGATGCTAACTAGTGTGGGCTTATTAGCTTGTGAAGAACTAAACCGCTGA 840  
Qy 841 AGTTGACGCGCTGGGAGATACGCTGCAAGATTAAACTCAAGAAATTGAAGGAGACCC 900  
Db 841 AGTTGACGCGCTGGGAGATACGCTGCAAGATTAAACTCAAGAAATTGAAGGAGACCC 900  
Qy 901 GCACAAAGCGGTGATATGATGATTATTGATGCAACGCGAAAAAACCCTTACCTT 960  
Db 901 GCACAAAGCGGTGATATGATGATTATTGATGCAACGCGAAAAAACCCTTACCTT 960  
Qy 961 GACATGTAGCCAAATTTTCTAGAGATAGATTAGTCTCGGGAACGCTAACAGAGTGTG 1020  
Db 961 GACATGTAGCCAAATTTTCTAGAGATAGATTAGTCTCGGGAACGCTAACAGAGTGTG 1020  
Qy 1021 CATGCTGTGCTGACGCTCGTGTGATGATGTTGGGTTAAGTCCGCAACGAGCCAA 1080  
Db 1021 CATGCTGTGCTGACGCTCGTGTGATGATGTTGGGTTAAGTCCGCAACGAGCCAA 1080  
Qy 1081 CTGTGCTAATTAATGCTCAATTTGTTGGGCACTTTAATGAGATGCGCGTGAACAA 1140  
Db 1081 CTGTGCTAATTAATGCTCAATTTGTTGGGCACTTTAATGAGATGCGCGTGAACAA 1140  
Qy 1141 GAGGAAGGTGGGAGATGACGTCAGTCCATAGGCGCTTATGAGTGGGTTTACAGTAA 1200  
Db 1141 GAGGAAGGTGGGAGATGACGTCAGTCCATAGGCGCTTATGAGTGGGTTTACAGTAA 1200  
Qy 1201 TACATGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTATCTCAGAAAGCGG 1260  
Db 1201 TACATGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTATCTCAGAAAGCGG 1260  
Qy 1261 TCGTAGTCGGATCGGAGTCTGCACTCGACTCCGTGAATCGGAATCGTAAATCGC 1320  
Db 1261 TCGTAGTCGGATCGGAGTCTGCACTCGACTCCGTGAATCGGAATCGTAAATCGC 1320  
Qy 1321 GATTCAGATGTCGGGAGTGAATCGTCCCGGCTTGTGACACACCGCCGTCACACAT 1380  
Db 1321 GATTCAGATGTCGGGAGTGAATCGTCCCGGCTTGTGACACACCGCCGTCACACAT 1380  
Qy 1381 GGGAGTGGGTTTCAACAGAGAGGTAGTCTAAACGTAAGAGAGGCGCTTGCACGTTGA 1440  
Db 1381 GGGAGTGGGTTTCAACAGAGAGGTAGTCTAAACGTAAGAGAGGCGCTTGCACGTTGA 1440  
Qy 1441 GATTTCATGACTGGGGTG 1457  
Db 1441 GATTTCATGACTGGGGTG 1457

RESULT 6  
AX316093 1457 bp DNA linear PAT 14-DEC-2001  
LOCUS AX316093  
DEFINITION Sequence 2 from Patent WO0190312.  
ACCESSION AX316093  
VERSION AX316093.1 GI:17899284  
KEYWORDS  
ORGANISM  
SOURCE  
unidentified  
unclassified  
unclassified sequences.

REFERENCE  
1  
AUTHORS Hovanec, T. A. and Burrell, P. C.  
TITLE Ammonia-oxidizing bacteria  
JOURNAL Patent: WO 0190312-A 2 29-NOV-2001;  
AQUARIA, INC. (US)  
FEATURES  
Source location/Qualifiers  
1. 1457

ORIGIN  
Query Match 100.0%; Score 1457; DB 6; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 8.3e-16;  
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3264"  
/note="AOB Type A1 ammonia-oxidizing bacterium represented  
by R7 clone187"

Qy 1 ATTGAAACGCTGGCGCATGCTTTACACATGCAAGTCGAACGCGACGACGATGCTTGCAT 60  
Db 1 ATTGAAACGCTGGCGCATGCTTTACACATGCAAGTCGAACGCGACGACGATGCTTGCAT 60  
Qy 61 CTGGTGGAGTGGCGGACGCGGTAGTATGCAATCGCAAGTATTCAGAAAGGGGGTGA 120  
Db 61 CTGGTGGAGTGGCGGACGCGGTAGTATGCAATCGCAAGTATTCAGAAAGGGGGTGA 120  
Qy 121 ACGCATGGAAGATGCTAATACCGCATATCTTAAGAGAGAAAGCAAGGAGATCGAAA 180  
Db 121 ACGCATGGAAGATGCTAATACCGCATATCTTAAGAGAGAAAGCAAGGAGATCGAAA 180  
Qy 181 GACCTTGCGCTTTTGAAGCGCGCATGCTGATTTAGCTAGTTGGTAAAGCCCTAC 240  
Db 181 GACCTTGCGCTTTTGAAGCGCGCATGCTGATTTAGCTAGTTGGTAAAGCCCTAC 240  
Qy 241 CAAGCGCAGATCACTAGTATGCTGAGAGAGCAACAGCCACTGCGACTGAGACAG 300  
Db 241 CAAGCGCAGATCACTAGTATGCTGAGAGAGCAACAGCCACTGCGACTGAGACAG 300  
Qy 301 GCCAGACTCTTACGGGAGCGAGAGTGGGGAATTTTGAACATGGGCGCAAGCTGATC 360  
Db 301 GCCAGACTCTTACGGGAGCGAGAGTGGGGAATTTTGAACATGGGCGCAAGCTGATC 360  
Qy 361 CAGCAATGCGCGGTAGTGAAGAAAGGCTTGGGTTGAAAGCTCTTCACTGAGAAAGA 420  
Db 361 CAGCAATGCGCGGTAGTGAAGAAAGGCTTGGGTTGAAAGCTCTTCACTGAGAAAGA 420  
Qy 421 AAAGGTTACGGTAAATTAATCTGTACTGACGCTATTCAGAAAGAACCGGCTTAAC 480  
Db 421 AAAGGTTACGGTAAATTAATCTGTACTGACGCTATTCAGAAAGAACCGGCTTAAC 480  
Qy 481 TACGTCGAGAGGCGCGGTAAATCGTAGGGTGAAGGCTTAATCGGAATTAAGTGGGCT 540  
Db 481 TACGTCGAGAGGCGCGGTAAATCGTAGGGTGAAGGCTTAATCGGAATTAAGTGGGCT 540  
Qy 541 AAAGGTCGCGAGCGCGCTTTGTAAGTCAGATGTGAATTCGCCGCTTAACTTGGGAT 600  
Db 541 AAAGGTCGCGAGCGCGCTTTGTAAGTCAGATGTGAATTCGCCGCTTAACTTGGGAT 600  
Qy 601 TCGCTTTGAAACTCAAGGCTAGAGTGTGCGAGAGGAGTGAATTCAGATGTAGAG 660  
Db 601 TCGCTTTGAAACTCAAGGCTAGAGTGTGCGAGAGGAGTGAATTCAGATGTAGAG 660  
Qy 661 TGAATGCGTGAAGATATGAAAGAAATCGATGCGAGAGGAGGCTCTGGGTTAAACT 720  
Db 661 TGAATGCGTGAAGATATGAAAGAAATCGATGCGAGAGGAGGCTCTGGGTTAAACT 720  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTTGGTAGTCCAGCC 780  
Db 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTTGGTAGTCCAGCC 780  
Qy 781 CTAAACGATGCTAACTAGTGTGGGCTTATTAGCTTGTGAAGAACTAAACCGCTGA 840  
Db 781 CTAAACGATGCTAACTAGTGTGGGCTTATTAGCTTGTGAAGAACTAAACCGCTGA 840  
Qy 841 AGTTGACGCGCTGGGAGATACGCTGCAAGATTAAACTCAAGAAATTGAAGGAGACCC 900  
Db 841 AGTTGACGCGCTGGGAGATACGCTGCAAGATTAAACTCAAGAAATTGAAGGAGACCC 900  
Qy 901 GCACAAAGCGGTGATATGATGATTATTGATGCAACGCGAAAAAACCCTTACCTT 960



Db 901 GCACAGCGGTGATTAATGATTAATTCGATGCAACCGGAAAACTTACCTACCTT 960  
Qy 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTCTTCGGGAACGCTAACAGAGTGTG 1020  
Db 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTCTTCGGGAACGCTAACAGAGTGTG 1020  
Qy 1021 CATGCGTGTGCTCAGCTGTGTGTGAGATGTTGGTTAAGTCCCGCAACGACCGCAC 1080  
Db 1021 CATGCGTGTGCTCAGCTGTGTGTGAGATGTTGGTTAAGTCCCGCAACGACCGCAC 1080  
Qy 1081 CTTGTGATTAATTTGCTATTTGTTGGGCACTTTAATGAGACTGCGGATGCAAAACG 1140  
Db 1081 CTTGTGATTAATTTGCTATTTGTTGGGCACTTTAATGAGACTGCGGATGCAAAACG 1140  
Qy 1141 GAGGAGGTGGGATGACGTCAGATCTCTATGAGCCCTTATGGGTAGGGCTTACACGTA 1200  
Db 1141 GAGGAGGTGGGATGACGTCAGATCTCTATGAGCCCTTATGGGTAGGGCTTACACGTA 1200  
Qy 1201 TACAAATGGCGCTACAGAGGTTGCCAACCCGCGAGGGAGCTAATCTCAGAAACGCG 1260  
Db 1201 TACAAATGGCGCTACAGAGGTTGCCAACCCGCGAGGGAGCTAATCTCAGAAACGCG 1260  
Qy 1261 TCGTACTCCGATCGGATGCTGCACTCGACTCCGTAAGTCCGTAATGCTTAATCCG 1320  
Db 1261 TCGTACTCCGATCGGATGCTGCACTCGACTCCGTAAGTCCGTAATGCTTAATCCG 1320  
Qy 1321 GGATGAGCATGTCGGGCTGATACGTTCCGGGCTTTGTACACACCGCCGCTACACAT 1380  
Db 1321 GGATGAGCATGTCGGGCTGATACGTTCCGGGCTTTGTACACACCGCCGCTACACAT 1380  
Qy 1381 GGGAGTGGGTTTCCAGCAGAGCAGTATCTAACTTAAGAGAGGGCGCTTCCACGCTGA 1440  
Db 1381 GGGAGTGGGTTTCCAGCAGAGCAGTATCTAACTTAAGAGAGGGCGCTTCCACGCTGA 1440  
Qy 1441 GATTCAATGACTGGGCTG 1457  
Db 1441 GATTCAATGACTGGGCTG 1457

RESULT 7  
AF386752 1457 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosomonas sp. R7c131 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386752  
ACCESSION AF386752.1 GI:17864829  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. R7c131  
ORGANISM Nitrosomonas sp. R7c131  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas; environmental samples.  
REFERENCE 1 (bases 1 to 1457)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 11722936  
REFERENCES 2 (bases 1 to 1457)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condon Dr, Moorpark, CA 93021, USA  
FEATURES  
source location/Qualifiers  
1..1457  
/organism="Nitrosomonas sp. R7c131"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichments"  
/db\_xref="taxon:180511"  
/clone="R7c131"  
/environmental\_sample  
<1..>1457  
/product="16S ribosomal RNA"

ORIGIN

rRNA

<1..>1457  
/product="16S ribosomal RNA"

Query Match 99.9%; Score 1455.4; DB 3; Length 1457;  
Best Local Similarity 99.9%; Pred. No. 8.8e-16;  
Matches 1456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATGAAAGCGGCGCATGCTTTTACATGCAAGTCGAACCGCACGAGATCTTGCAT 60  
Db 1 ATGAAAGCGTGGCGCATGCTTTTACATGCAAGTCGAACCGCACGAGATCTTGCAT 60  
Qy 61 CTGTGTGGAGTGGCGGACGCGGTGATATGATGGAACGTATCCAGAAAGGGGGGTA 120  
Db 61 CTGTGTGGAGTGGCGGACGCGGTGATATGATGGAACGTATCCAGAAAGGGGGGTA 120  
Qy 121 ACCGATCCAGAAATGTGCTTAATCCGCGATATCTCTAAGAGAGAAACGGGGATCGAA 180  
Db 121 ACCGATCCAGAAATGTGCTTAATCCGCGATATCTCTAAGAGAGAAACGGGGATCGAA 180  
Qy 181 GACCTTGGCGCTTTTGAAGCGCGCATGTCGATTAAGTGTGTTGGGTAAAGGCTTAC 240  
Db 181 GACCTTGGCGCTTTTGAAGCGCGCATGTCGATTAAGTGTGTTGGGTAAAGGCTTAC 240  
Qy 241 CAAGGCGAGATCAGTATGTTGTGAGAGACGACCACTGGAATGAGACAG 300  
Db 241 CAAGGCGAGATCAGTATGTTGTGAGAGACGACCACTGGAATGAGACAG 300  
Qy 301 GCCCAGACTCTTACCGGAGGACGAGTGGGAAATTTTGAACATGGGCGCAAGCCTGATC 360  
Db 301 GCCCAGACTCTTACCGGAGGACGAGTGGGAAATTTTGAACATGGGCGCAAGCCTGATC 360  
Qy 361 CAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAGA 420  
Db 361 CAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAGA 420  
Qy 421 AAAAGTTACGTTAATTAATCTGACTCATGACGATTCGACAGAAAGACACCGGCTAAC 480  
Db 421 AAAAGTTACGTTAATTAATCTGACTCATGACGATTCGACAGAAAGACACCGGCTAAC 480  
Qy 481 TACGTCCAGACAGCCCGGTAAATACGTAGGGTGCAGGCGTTAATCGAAATTAATGAGCG 540  
Db 481 TACGTCCAGACAGCCCGGTAAATACGTAGGGTGCAGGCGTTAATCGAAATTAATGAGCG 540  
Qy 541 AAAAGGTGGCGAGGGCGCTTTGTAAGTCAGATGTGAATCCCGGCTTAACTGGGAAT 600  
Db 541 AAAAGGTGGCGAGGGCGCTTTGTAAGTCAGATGTGAATCCCGGCTTAACTGGGAAT 600  
Qy 601 TCGCTTTGAAACTACAGGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTACAG 660  
Db 601 TCGCTTTGAAACTACAGGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTACAG 660  
Qy 661 TGAATGCGTAGATATGAGAGACATGATGCGAAGGACGCTCCCTGGTTAATCACT 720  
Db 661 TGAATGCGTAGATATGAGAGACATGATGCGAAGGACGCTCCCTGGTTAATCACT 720  
Qy 721 GACGCTCAGCAGAAAGGTGGGAGCAAAACGAGTTAGTACCCTGTGATCCAGGCC 780  
Db 721 GACGCTCAGCAGAAAGGTGGGAGCAAAACGAGTTAGTACCCTGTGATCCAGGCC 780  
Qy 781 CTAACGATGCAACTAGTGTGCTTATGAGCTTATGAGCTTGTGAACGAACTAACGCTGA 840  
Db 781 CTAACGATGCAACTAGTGTGCTTATGAGCTTATGAGCTTGTGAACGAACTAACGCTGA 840  
Qy 841 AGTTGACCGCTGGGGGATACGCTGCCAATTTAAATCTAAAGAAATTAACGGGAGCC 900  
Db 841 AGTTGACCGCTGGGGGATACGCTGCCAATTTAAATCTAAAGAAATTAACGGGAGCC 900  
Qy 901 GCACAGCGGTGATATGATGATTAATGATGCAACGGAAGAACTTACCTACCTT 960  
Db 901 GCACAGCGGTGATATGATGATTAATGATGCAACGGAAGAACTTACCTACCTT 960  
Qy 961 GACATGTAGCAATTTTCTAGAGATAGATTAGTGTCCGGAACGCTAACACAGGTGCTG 1020  
Db 961 GACATGTAGCAATTTTCTAGAGATAGATTAGTGTCCGGAACGCTAACACAGGTGCTG 1020

QY 1021 CATGGCTGTCAGCTGTCGTGAGATGTTGGTTAAAGTCCCGCAACGACCGCAACC 1080  
DB 1021 CATGGCTGTCAGCTGTCGTGAGATGTTGGTTAAAGTCCCGCAACGACCGCAACC 1080  
QY 1081 CTTGTCAATTAATTTGGCATATTTGGTTGGGCACTTTAATAGATCTGCCGTGACAAACCG 1140  
DB 1081 CTTGTCAATTAATTTGGCATATTTGGTTGGGCACTTTAATAGATCTGCCGTGACAAACCG 1140  
QY 1141 GAGGAAGGTGGGAGATGACGTCAAGTCTCATGSCCTTTATGGGTAGGGCTTTACACGTAA 1200  
DB 1141 GAGGAAGGTGGGAGATGACGTCAAGTCTCATGSCCTTTATGGGTAGGGCTTTACACGTAA 1200  
QY 1201 TACAAATGGCGGTACAGAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAAGCCGG 1260  
DB 1201 TACAAATGGCGGTACAGAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAAGCCGG 1260  
QY 1261 TCGTAGTCCGAGTCGGAAGTCTGCAACTCGACTCCGTGAAGTCGGAATCGCTAGTAATCGC 1320  
DB 1261 TCGTAGTCCGAGTCGGAAGTCTGCAACTCGACTCCGTGAAGTCGGAATCGCTAGTAATCGC 1320  
QY 1321 GGATCAGCATGTTCGGGTGATTAACGTTCCGGGCTTTGTACACACCGCCCTGACACCAT 1380  
DB 1321 GGATCAGCATGTTCGGGTGATTAACGTTCCGGGCTTTGTACACACCGCCCTGACACCAT 1380  
QY 1381 GGGAGTGGGTTTACACAGAAAGAGGTAGTCTAACCGTAAAGAGGGGCGCTTGCACAGGTGA 1440  
DB 1381 GGGAGTGGGTTTACACAGAAAGAGGTAGTCTAACCGTAAAGAGGGGCGCTTGCACAGGTGA 1440  
QY 1441 GATTTCATGACTGGGGTG 1457  
DB 1441 GATTTCATGACTGGGGTG 1457  
RESULT 8  
CQ796909 1457 bp DNA linear PAT 19-APR-2004  
LOCUS CQ796909 Sequence 2 from Patent WO2004026772.  
DEFINITION CQ796909  
ACCESSION CQ796909.1 GI:46408535  
VERSION CQ796909.1 GI:46408535  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1  
AUTHORS Hovanec, T. A.  
TITLE Ammonite-oxidizing bacteria and methods of using and detecting  
JOURNAL chesame  
Patent: WO 2004026772-A 2 01-APR-2004;  
Aquaaria Inc. (US)  
FEATURES  
Source location/Qualifiers  
1..1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="AOB Type A1 R7clone187 16S rDNA"  
ORIGIN  
Query Match 99.7%; Score 1452.2; DB 6; Length 1457;  
Beet Local Similarity 99.8%; Pred. No. 9.8e-16;  
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ATTGAACGCTGGCGGAGATCTTTACACATGCAAGTCGAAACGGGACACAGAGATCTTGAT 60  
DB 1 ATTGAACGCTGGCGGAGATCTTTACACATGCAAGTCGAAACGGGACACAGAGATCTTGAT 60  
QY 61 CTGGTGGCGAGTGGCGGAGCGGGTGAATATGATCGAAGCGTATCCAGAAAGAGGGGGTA 120  
DB 61 CTGGTGGCGAGTGGCGGAGCGGGTGAATATGATCGAAGCGTATCCAGAAAGAGGGGGTA 120  
QY 121 ACGCATCGAAAGATGTCTAATATCCGATATATCTTAAAGAGAAAGCAGGGGATCGAA 180  
DB 121 ACGCATCGAAAGATGTCTAATATCCGATATATCTTAAAGAGAAAGCAGGGGATCGAA 180

QY 181 GACCTTGCGCTTTTGGAGCGGCCGATGTCTGATAGTAGTGGTGGGCTTAAAGGCTTAC 240  
DB 181 GACCTTGCGCTTTTGGAGCGGCCGATGTCTGATAGTAGTGGTGGGCTTAAAGGCTTAC 240  
QY 241 CAAGGCGACGATCAAGTATGTTGGTCTGAGAGAGCAGCAGCAGCATCTGGAACTAGACACG 300  
DB 241 CAAGGCGACGATCAAGTATGTTGGTCTGAGAGAGCAGCAGCAGCATCTGGAACTAGACACG 300  
QY 301 GCCCAGATCTCTACGGGAGGAGCAGAGTGGGGAATTTTGGACAAATGGGGGCAAGCTGATC 360  
DB 301 GCCCAGATCTCTACGGGAGGAGCAGAGTGGGGAATTTTGGACAAATGGGGGCAAGCTGATC 360  
QY 361 CAGCAATCCCGGTAGAGTGAAGAAAGGCTTGGGGTGTAAAGCTTTTCAAGTCGAGAAAGA 420  
DB 361 CAGCAATCCCGGTAGAGTGAAGAAAGGCTTGGGGTGTAAAGCTTTTCAAGTCGAGAAAGA 420  
QY 421 AAAGTTACGGTAAATATATCGTGAATCTCATGACGGTATCGACAGAAAGACACCGGCTAAC 480  
DB 421 AAAGTTACGGTAAATATATCGTGAATCTCATGACGGTATCGACAGAAAGACACCGGCTAAC 480  
QY 481 TACGTCCAGCAGCGCGCGGTAAATACGTAGGGGTGCAAGGCTTAATCGGAATTAATCTGGGCGT 540  
DB 481 TACGTCCAGCAGCGCGCGGTAAATACGTAGGGGTGCAAGGCTTAATCGGAATTAATCTGGGCGT 540  
QY 541 AAAGGTCGCGAGCGGCTTTGTAAAGTGAATGTAATCCCGGGCTTAACTGGGAAT 600  
DB 541 AAAGGTCGCGAGCGGCTTTGTAAAGTGAATGTAATCCCGGGCTTAACTGGGAAT 600  
QY 601 TCGGTTGAAACTACAAAGCTAGAGTGTGAGAGAGGAGGTGAATTCATGTGTAGAG 660  
DB 601 TCGGTTGAAACTACAAAGCTAGAGTGTGAGAGAGGAGGTGAATTCATGTGTAGAGAG 660  
QY 661 TGAATGTGTAGATATGAGAAACATTCATGCGCAAGGCGCTCTGGGTTAAACCT 720  
DB 661 TGAATGTGTAGATATGAGAAACATTCATGCGCAAGGCGCTCTGGGTTAAACCT 720  
QY 721 GACGCTCATGACGAAAGCGTGGGAGAGCAAGATTAATACCTGTAGTCCACGCC 780  
DB 721 GACGCTCATGACGAAAGCGTGGGAGAGCAAGATTAATACCTGTAGTCCACGCC 780  
QY 781 CTAAACGATGTCATAGTGTGGGCTTTATAGGCTTGGTAAAGAACTAACGCGGTGA 840  
DB 781 CTAAACGATGTCATAGTGTGGGCTTTATAGGCTTGGTAAAGAACTAACGCGGTGA 840  
QY 841 AGTTGACCGCTGGGAGTACGCTGCGCAAGTTTAAACTCAAGAAATGACGGGAGACC 900  
DB 841 AGTTGACCGCTGGGAGTACGCTGCGCAAGTTTAAACTCAAGAAATGACGGGAGACC 900  
QY 901 GCACAAAGCGGTGATATATGATGATTAATTCGATGCAACGCGAAAAACCTTACCTTACCCTT 960  
DB 901 GCACAAAGCGGTGATATATGATGATTAATTCGATGCAACGCGAAAAACCTTACCTTACCCTT 960  
QY 961 GACATGTAGCAATTTTCTAGAGATAGATTAGTCTTGGGAAACGTTAACACAGGTGCTG 1020  
DB 961 GACATGTAGCAATTTTCTAGAGATAGATTAGTCTTGGGAAACGTTAACACAGGTGCTG 1020  
QY 1021 CATGGCTGTGCTCAGCTGCTGTCGTGAGATGTTGGGTTAAAGTCCCGCAACGAGGCAACC 1080  
DB 1021 CATGGCTGTGCTCAGCTGCTGTCGTGAGATGTTGGGTTAAAGTCCCGCAACGAGGCAACC 1080  
QY 1081 CTTGTCAATTAATTTGGCATATTTGGTTGGGCACTTTAATAGATCTGCCGTGACAAACCG 1140  
DB 1081 CTTGTCAATTAATTTGGCATATTTGGTTGGGCACTTTAATAGATCTGCCGTGACAAACCG 1140  
QY 1141 GAGGAAGGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACGTAA 1200  
DB 1141 GAGGAAGGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACGTAA 1200  
QY 1201 TACAAATGGCGGTACAGAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
DB 1201 TACAAATGGCGGTACAGAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
QY 1261 TCGTAGTCCGAGTCGGAAGTCTGCAACTCGACTCCGTGAAGTCGGAATCGCTAATATCGC 1320

|||||  
Db 1261 TCGATCGGATCGGAGTCTCAACTCCATCCGTAATCGGATCGCTGTATCGC 1320  
|||  
Qy 1321 GGATAGCATGTCGGGTGAATACGTTCCGGGTCTTTGACACCGCCGCTCACACAT 1380  
|||  
Db 1321 GGATAGCATGTCGGGTGAATACGTTCCGGGTCTTTGACACCGCCGCTCACACAT 1380  
|||  
Qy 1381 GGGAGTGGGTTTACCAAGAGAGGATGTAACCTTAAGAGAGGCGCTTGGCCAGGTGA 1440  
|||  
Db 1381 GGGAGTGGGTTTACCAAGAGAGGATGTAACCTTAAGAGAGGCGCTTGGCCAGGTGA 1440  
|||  
Qy 1441 GATTCATGACTGGGGTG 1457  
|||  
Db 1441 GATTCATGACTGGGGTG 1457  
|||  
RESULT 9  
AF386757 1457 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosomonas sp. R7c187 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386757  
ACCESSION AF386757.1 GI:117864834  
VERSION  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. R7c187  
ORGANISM Nitrosomonas sp. R7c187  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas; environmental samples.  
REFERENCE 1 (bases 1 to 1457)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 11722936  
2 (bases 1 to 1457)  
REFERENCE Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
AUTHORS Direct Submision  
TITLE Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
JOURNAL Group, 6100 Condon Dr., Moorpark, CA 93021, USA  
FEATURES  
source  
1. 1457  
/organism="Nitrosomonas sp. R7c187"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichment"  
/db\_xref="taxon:180512"  
/clone="R7c187"  
/environmental\_sample  
<1. >1457  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 99.5%; Score 1449; DB 3; Length 1457;  
Best Local Similarity 99.7%; Pred. No. 1.1e-15;  
Matches 1452; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 ATTTGAACGCTGGCGGACGCTTTTACATGCAATCGAAGCGGACGACGATGCTTGCAAT 60  
Db 1 ATTTGAACGCTGGCGGACGCTTTTACATGCAATCGAAGCGGACGACGATGCTTGCAAT 60  
|||  
Qy 61 CTGGTGGCGAGTGGCGGACGCGGTGATGATGCAATCGGAACGTATCCGAAGAGGGGGTGA 120  
Db 61 CTGGTGGCGAGTGGCGGACGCGGTGATGATGCAATCGGAACGTATCCGAAGAGGGGGTGA 120  
|||  
Qy 121 ACCGATGGAAGATGTGCTAATACCGCATATACCTTAAGAGGAAGACAGGGGATCGAAA 180  
Db 121 ACCGATGGAAGATGTGCTAATACCGCATATACCTTAAGAGGAAGACAGGGGATCGAAA 180  
|||  
Qy 181 GACCTTGGCTTTTGAAGCGGCGGATGCTGATTAAGCTAAGTGGTGGGATAAAGCCGTAC 240  
Db 181 GACCTTGGCTTTTGAAGCGGCGGATGCTGATTAAGCTAAGTGGTGGGATAAAGCCGTAC 240  
|||  
Qy 241 CAAGCGACGATCAGTAGTGGTCTGAGAGGACGACCACTGGGACTGAGACAG 300  
Db 241 CAAGCGACGATCAGTAGTGGTCTGAGAGGACGACCACTGGGACTGAGACAG 300  
|||

Db 241 CAAGCGACGATCAGTAGTGGTCTGAGAGGACGACCACTGGGACTGAGACAG 300  
|||  
Qy 301 GCCCAAGCTCCTTACGGGAGGCGAGCGTGGGAAATTTTGAAGAAATGGGGGCAAGCCTGATC 360  
|||  
Db 301 GCCCAAGCTCCTTACGGGAGGCGAGCGTGGGAAATTTTGAAGAAATGGGGGCAAGCCTGATC 360  
|||  
Qy 361 CAGCAATCCCGGTGAGGAGGAAGAGCCCTCGGGTTTAAAGCTCTTTCACTGAGAGAGA 420  
|||  
Db 361 CAGCAATCCCGGTGAGGAGGAAGAGCCCTCGGGTTTAAAGCTCTTTCACTGAGAGAGA 420  
|||  
Qy 421 AAAGGTTACGGTAATTAATCGTACTGATGACGATGCGAGAGAGAACACCGGCTAAC 480  
Db 421 AAAGGTTACGGTAATTAATCGTACTGATGACGATGCGAGAGAGAACACCGGCTAAC 480  
|||  
Qy 481 TACGTCCAGACAGCCCGGTAATACGTAGGGTGCAGAGCTTAAATCGGAATTAATCGGGCGT 540  
Db 481 TACGTCCAGACAGCCCGGTAATACGTAGGGTGCAGAGCTTAAATCGGAATTAATCGGGCGT 540  
|||  
Qy 541 AAAGGTTGCGGAGCGGCGCTTTGTAAGTACGATGTGAATCCCGGGGCTTAACCTGGGAAT 600  
Db 541 AAAGGTTGCGGAGCGGCGCTTTGTAAGTACGATGTGAATCCCGGGGCTTAACCTGGGAAT 600  
|||  
Qy 601 TCGCTTTGAATCTAACAAGCTAGAGTGGGACAGAGGAGTGAATTCATGTGTAGCAG 660  
Db 601 TCGCTTTGAATCTAACAAGCTAGAGTGGGACAGAGGAGTGAATTCATGTGTAGCAG 660  
|||  
Qy 661 TGAATGCGTAGAGATATGAAGAAACATGATGCGGAGGCGCTCTCTGGGTTAACT 720  
Db 661 TGAATGCGTAGAGATATGAAGAAACATGATGCGGAGGCGCTCTCTGGGTTAACT 720  
|||  
Qy 721 GACGCTATGACGAAAGCGTGGGAGCAACAGAGTTAGATACCTGTAGTCCAGCC 780  
Db 721 GACGCTATGACGAAAGCGTGGGAGCAACAGAGTTAGATACCTGTAGTCCAGCC 780  
|||  
Qy 781 CTAAACGATGTCAACTAGTTGTGGGCTTAATTAAGCTTGTATACGAACTAACCGGTGA 840  
Db 781 CTAAACGATGTCAACTAGTTGTGGGCTTAATTAAGCTTGTATACGAACTAACCGGTGA 840  
|||  
Qy 841 AGTTGACGCGCTGGGAGTACCGGTGCGAAGATTAATACTCAAGAAATTGACGGGAGACC 900  
Db 841 AGTTGACGCGCTGGGAGTACCGGTGCGAAGATTAATACTCAAGAAATTGACGGGAGACC 900  
|||  
Qy 901 GCAACAAGCGGTGATATATGATGATTAATTCATGCAACGCGAATACTTACCTACCTT 960  
Db 901 GCAACAAGCGGTGATATATGATGATTAATTCATGCAACGCGAATACTTACCTACCTT 960  
|||  
Qy 961 GACATGAGGAAATTTTCTAGAGATAGATTAGTCTTGGGAACTGAAACAGAGTGTCTG 1020  
Db 961 GACATGAGGAAATTTTCTAGAGATAGATTAGTCTTGGGAACTGAAACAGAGTGTCTG 1020  
|||  
Qy 1021 CATGCTGTCTGACGCTGCTGTCGAGATGTTGGTTAAATGTCGCCAACGAGCGCAACC 1080  
Db 1021 CATGCTGTCTGACGCTGCTGTCGAGATGTTGGTTAAATGTCGCCAACGAGCGCAACC 1080  
|||  
Qy 1081 CTGTGATTAATGTCATCATTTTGTGGGCACTTTAAATGAGCTGCGGTGCAAAACCG 1140  
Db 1081 CTGTGATTAATGTCATCATTTTGTGGGCACTTTAAATGAGCTGCGGTGCAAAACCG 1140  
|||  
Qy 1141 GAGGAAGTGGGAGTACGTAAGTCTCATGAGCTTATGAGGAGGCTTACACAGTAA 1200  
Db 1141 GAGGAAGTGGGAGTACGTAAGTCTCATGAGCTTATGAGGAGGCTTACACAGTAA 1200  
|||  
Qy 1201 TACAAATGGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTTAATCTCAAGAAAGCGC 1260  
Db 1201 TACAAATGGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTTAATCTCAAGAAAGCGC 1260  
|||  
Qy 1261 TCGTAGTCCGAGTCCGAGTGTGCACTGCACTCGTGAAGTCCGAATCGTAAATTCGC 1320  
Db 1261 TCGTAGTCCGAGTCCGAGTGTGCACTGCACTCGTGAAGTCCGAATCGTAAATTCGC 1320  
|||  
Qy 1321 GGATCAGCATGTGCGGGTGAATACGTTCCGGGCTTTGTATACACCGCCGCTCACACAT 1380  
Db 1321 GGATCAGCATGTGCGGGTGAATACGTTCCGGGCTTTGTATACACCGCCGCTCACACAT 1380  
|||

QY 1381 GGGAGTGGTTTCCACGAAGAAGAGTACTTAACCGTAGAGAGGGCGCTTGGCAGGTGA 1440  
Db 1381 GGGAGTGGTTTCCACGAAGAAGAGTACTTAACCGTAGAGAGGGCGCTTGGCAGGTGA 1440  
QY 1441 GATTGATGACTGGGGTG 1457  
Db 1441 GATTGATGACTGGGGTG 1457

RESULT 10  
AF386751 1426 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosomonas sp. R7c155 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386751.1 GI:17864828  
VERSION AF386751.1 GI:17864828  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. R7c155  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas; environmental samples.  
REFERENCE 1 (bases 1 to 1426)  
AUTHORS Burrell,P.C., Phalen,C.M. and Hovanec,T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 11722936  
REFERENCE 2 (bases 1 to 1426)  
AUTHORS Burrell,P.C., Phalen,C.M. and Hovanec,T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condy Dr., Moorpark, CA 93021, USA

FEATURES  
source  
1..1426  
/organism="Nitrosomonas sp. R7c155"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichment"  
/db\_xref="taxon:160510"  
/clone="R7c155"  
/environmental\_sample  
<1..>1426  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 97.3%; Score 1417; DB 3; Length 1426;  
Best Local Similarity 99.6%; Pred. No. 3.6e-15;  
Matches 1420; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 32 AAGTGAACGGCAGACGGATGCTTGCACTGTGGCGAGTGGCGGAGCGGTGAGTATG 91  
Db 1 AAGTGAACGGCAGACGGATGCTTGCACTGTGGCGAGTGGCGGAGCGGTGAGTATG 60  
QY 92 CATCGGAACGTATCCAGAAAGAGGGGGGTAAACGATCGAAAGATGTCTAATACCGCATAT 151  
Db 61 CATCGGAACGTATCCAGAAAGAGGGGGGTAAACGATCGAAAGATGTCTAATACCGCATAT 120  
QY 152 ACTTGAAGAGAAAGCAGGGGATCGAAAGACCTTGGCGCTTTTGAAGCGGCCGATGTCTG 211  
Db 121 ACTTGAAGAGAAAGCAGGGGATCGAAAGACCTTGGCGCTTTTGAAGCGGCCGATGTCTG 180  
QY 212 ATTAGCTAGTGGTGGGGTAAAGGCGCTACCAAGGGAGAGATAGTATGTTGCTGAGAGG 271  
Db 181 ATTAGCTAGTGGTGGGGTAAAGGCGCTACCAAGGGAGAGATAGTATGTTGCTGAGAGG 240  
QY 272 ACGACCAACCACTGGGAGCTGAGACACGGCCAGACTTCTTACGGGAGGCAAGCAGTGGGG 331  
Db 241 ACGACCAACCACTGGGAGCTGAGACACGGCCAGACTTCTTACGGGAGGCAAGCAGTGGGG 300  
QY 332 AATTTGGAACAATGGGGCGCAAGCTGATCAGCAATGCCGCTGAGTGAAGAAAGCCTTC 391  
Db 301 AATTTGGAACAATGGGGCGCAAGCTGATCAGCAATGCCGCTGAGTGAAGAAAGCCTTC 360

QY 392 GGGTTGAAGCTCTTTCACTCGAAGAAAGAAAGTTACGTAATAATCGTACTCATGA 451  
Db 361 GGGTTGAAGCTCTTTCACTCGAAGAAAGAAAGTTACGTAATAATCGTACTCATGA 420  
QY 452 CGGTATCGACAGAAAGAACACCGGCTTAACTAGTGCACAGCCCGGGTAAATACGTAAGG 511  
Db 421 CGGTATCGACAGAAAGAACACCGGCTTAACTAGTGCACAGCCCGGGTAAATACGTAAGG 480  
QY 512 TGCAGCGTTAATCGGAATTAATCTGGCGGTAAAGGGTGGCAGGGCGCTTTGAAGTCA 571  
Db 481 TGCAGCGTTAATCGGAATTAATCTGGCGGTAAAGGGTGGCAGGGCGCTTTGAAGTCA 540  
QY 572 TGTGAATCCCGGGGCTTAACTGGGAATTCGTTTGAATCAACAGCCTAGAGTGGC 631  
Db 541 TGTGAATCCCGGGGCTTAACTGGGAATTCGTTTGAATCAACAGCCTAGAGTGGC 600  
QY 632 AGAGGAGTGAATTCATGTGTAGCAGTGAATAATGCTAGAGATATGAAGAAATCTGA 691  
Db 601 AGAGGAGTGAATTCATGTGTAGCAGTGAATAATGCTAGAGATATGAAGAAATCTGA 660  
QY 692 TGGCGAAGCGACCTCCCTGGGTTAACTGACGCTCATGACGAAAGCGTGGGAGCAAA 751  
Db 661 TGGCGAAGCGACCTCCCTGGGTTAACTGACGCTCATGACGAAAGCGTGGGAGCAAA 720  
QY 752 CAGGATTGATACCTCTGTAGTCAACGCGCTTAAAGATGTCACTAGTTGTTGGGCCCTTA 811  
Db 721 CAGGATTGATACCTCTGTAGTCAACGCGCTTAAAGATGTCACTAGTTGTTGGGCCCTTA 780  
QY 812 TTAGGCTTGTAAAGAGCTAAACGGGTGAAGTTGACCGGCTGGGAGATACGCTGCAGAA 871  
Db 781 TTAGGCTTGTAAAGAGCTAAACGGGTGAAGTTGACCGGCTGGGAGATACGCTGCAGAA 840  
QY 872 TTTAAACCTCAAGGAATTTGACGGGGACCCGCAAGGGGTGATTAATGTGATTAATTCG 931  
Db 841 TTTAAACCTCAAGGAATTTGACGGGGACCCGCAAGGGGTGATTAATGTGATTAATTCG 900  
QY 932 ATGCAACCGCAAAAACTTACCTTACCTTGAATGTAGCAATTTTCTAGAGATAGATTA 991  
Db 901 ATGCAACCGCAAAAACTTACCTTACCTTGAATGTAGCAATTTTCTAGAGATAGATTA 960  
QY 992 GTGCTTCGGAACGCTAACACAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1051  
Db 961 GTGCTTCGGAACGCTAACACAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
QY 1052 TTGGGTTAATCCCGCAACGAGCCGCAACCTTGTCAATTAATTCGATCATTTGGTGGGC 1111  
Db 1021 TTGGGTTAATCCCGCAACGAGCCGCAACCTTGTCAATTAATTCGATCATTTGGTGGGC 1080  
QY 1112 ACTTTAATGAGACTGCCGCTGACAAACCGAGAGAGTGGGATGACGTCAGTCTTCAT 1171  
Db 1081 ACTTTAATGAGACTGCCGCTGACAAACCGAGAGAGTGGGATGACGTCAGTCTTCAT 1140  
QY 1172 GGGCTTTATGGGTAGGGCTTCAACGTATTAATGATGCGGCTACAGAGGTTGCCAACCC 1231  
Db 1141 GGGCTTTATGGGTAGGGCTTCAACGTATTAATGATGCGGCTACAGAGGTTGCCAACCC 1200  
QY 1232 GCGAGGGGAGCTAATCTCAAGAAAGCGGTGTGTCGCGATCGGAATCGCAACTGCAC 1291  
Db 1201 GCGAGGGGAGCTAATCTCAAGAAAGCGGTGTGTCGCGATCGGAATCGCAACTGCAC 1260  
QY 1292 TCCGTGAAGTGGGAATCGCTAGTAAATCGCGAGATCAGCATGTCCGCGTGAATACGTTCCCG 1351  
Db 1261 TCCGTGAAGTGGGAATCGCTAGTAAATCGCGAGATCAGCATGTCCGCGTGAATACGTTCCCG 1320  
QY 1352 GGTCTTGTACACACCGCCGCTCAACATGGGAATGGGTTTACCAAGAAAGAGTATGCT 1411  
Db 1321 GGTCTTGTACACACCGCCGCTCAACATGGGAATGGGTTTCAACAGAAAGAGTATGCT 1380  
QY 1412 AACCGTAAAGAGGGCGCTTGCACAGGTGAAGTTATATGCTGGGGTG 1457  
Db 1381 AACCGTAAAGAGGGCGCTTGCACAGGTGAAGTTATATGCTGGGGTG 1426

RESULT 11  
AF386746 1487 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosomonas sp. BF16c57 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386746  
VERSION AF386746.1 GI:17864823  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. BF16c57  
ORGANISM Nitrosomonas sp. BF16c57  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas; environmental samples.  
REFERENCE 1 (bases 1 to 1487)  
Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 1172936  
2 (bases 1 to 1487)  
Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condon Dr. Moorpark, CA 93021, USA  
FEATURES  
source  
1. 1487  
/organism="Nitrosomonas sp. BF16c57"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichments"  
/db\_xref="taxon:180505"  
/clone="BF16c57"  
/environmental\_sample  
<1. >1487  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 95.5%; Score 1391.6; DB 3; Length 1487;  
Best Local Similarity 97.6%; Pred. No. 8.5e-15;  
Matches 142; Conservative 0; Mismatches 34; Indels 1; Gaps 1;  
QY 1 ATTGAAGCTGGCGGCGATGCTTTACATGCAAGTCGACGGGCGACGCGATGTTSCAT 60  
DB 18 ATTGAAGCTGGCGGCGGCTTTACATGCAAGTCGACGGGCGACGCGGCTGTTSCAC 77  
QY 61 CTGGTGCAGATGGCGGCGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 78 CTGGTGCAGATGGCGGCGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 137  
QY 121 ACCGATCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
DB 138 ACCGATCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 197  
QY 181 GACCTTGGCTTTTGGAGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
DB 198 GACCTTGGCTTTTGGAGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 257  
QY 241 CAAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 258 CAAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 317  
QY 301 GCCAGACTCTTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
DB 318 GCCAGACTCTTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 377  
QY 361 CAGCAATGCCCGCTGAGTGAAGAGGCTTGGGTTGAAGCTTTTCACTCGAGAGAA 420  
DB 378 CAGCAATGCCCGCTGAGTGAAGAGGCTTGGGTTGAAGCTTTTCACTCGAGAGAA 437  
QY 421 AAAGTTACGGTAAATATCTGACTCATGACGGATGACAGAGAGAGAGAGAGAGAGAGAGAG 480  
DB 438 AAAGTTACGGTAAATATCACTTATGACGGATGACAGAGAGAGAGAGAGAGAGAGAGAG 497  
QY 481 TACGTCCAGAGCGCGGCTAATACGTAAGGTGCAAGCTTAATCGAATTAATCTGGGCGT 540

DB 498 TACGTCCAGAGCGCGGCTAATACGTAAGGTGCAAGCTTAATCGAATTAATCTGGGCGT 557  
QY 541 AAAGTTACGGTAAATATCTGACTCATGACGGATGACAGAGAGAGAGAGAGAGAGAGAGAG 600  
DB 558 AAAGTTACGGTAAATATCACTTATGACGGATGACAGAGAGAGAGAGAGAGAGAGAGAG 617  
QY 601 TGGCGTTGAAATCAAGGCTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 618 TGGCGTTGAAATCAAGGCTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 677  
QY 661 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
DB 678 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737  
QY 721 GACGCTCATGACAGAAAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
DB 738 GACGCTCATGACAGAAAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 797  
QY 781 CTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 798 CTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857  
QY 841 AGTTGACCGCTGGGAGGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 858 AGTTGACCGCTGGGAGGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 917  
QY 901 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 918 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 977  
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019  
DB 978 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1037  
QY 1020 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079  
DB 1038 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1097  
QY 1080 CTTTGTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139  
DB 1098 CTTTGTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157  
QY 1140 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1199  
DB 1158 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1217  
QY 1200 ATACAAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259  
DB 1218 ATACAAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1277  
QY 1260 GTGCTAGTCCGATGCGAGTCTGCAACTGCACTCCGTAAGTCCGTAAGTCCGTAAGTCCG 1319  
DB 1278 GTGCTAGTCCGATGCGAGTCTGCAACTGCACTCCGTAAGTCCGTAAGTCCGTAAGTCCG 1337  
QY 1320 CGGATCAGCATGCTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379  
DB 1338 CGGATCAGCATGCTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1397  
QY 1380 TGGAGTGGGTTTCAACAGAGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1439  
DB 1398 TGGAGTGGGTTTCAACAGAGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
QY 1440 AGATTCAATGACTGGGCTG 1457  
DB 1458 AGATTCAATGACTGGGCTG 1475

RESULT 12  
CQ796927  
LOCUS CQ796927 1491 bp DNA linear PAT 19-APR-2004  
DEFINITION Sequence 20 from Patent WO2004026772.  
ACCESSION CQ796927.1 GI:46408553  
VERSION

KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE unclassified sequences.  
AUTHORS 1  
TITLE Hovanec, T. A.  
JOURNAL Ammonia-oxidizing bacteria and methods of using and detecting  
the same  
Patent: WO 2004026772-A 20 01-APR-2004;  
Aquaeria Inc. (US)  
FEATURES Location/Qualifiers  
source 1..1491  
/organism="unclassified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="N. aestuarii-like AOB Bflc1ones57 16S rDNA"

ORIGIN  
Query Match 95.5%; Score 1391.6; DB 6; Length 1491;  
Best Local Similarity 97.6%; Pred. No. 8.5e-15;  
Matches 1423; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 1 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAACGGCAGCAGCGATGCTTGCAAT 60  
DB 18 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAACGGCAGCAGCGATGCTTGCAAT 77  
QY 61 CTGTGGCGAGTGGCGGAGCGGCTGAGTAATGCATCGGAACGTAATCGAAGAGGGGGGTA 120  
DB 78 CTGTGGCGAGTGGCGGAGCGGCTGAGTAATGCATCGGAACGTAATCGAAGAGGGGGGTA 137  
QY 121 AGCGATCGGAAGATGTGCTAATACCGCATATCTCTAAGAGGAAGAGGGGGATCGAAT 180  
DB 138 AGCGATCGGAAGATGTGCTAATACCGCATATCTCTAAGAGGAAGAGGGGGATCGAAT 197  
QY 181 GACCTTGCGCTTTTGGAGCGCGCGATGTCTGATTAAGCTAGTTGTGGGGTAAAGGCTTAC 240  
DB 198 GACCTTGCTTTTGGAGCGCGCGATGTCTGATTAAGCTAGTTGTGGGGTAAAGGCTTAC 257  
QY 241 CAAGGCGAGATCATAGTGTGCTGAGAGAGACACACGACCTGGGATCGAAGACAG 300  
DB 258 CAAGGCGAGATCATAGTGTGCTGAGAGAGACACACGACCTGGGATCGAAGACAG 317  
QY 301 GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCCCTGATC 360  
DB 318 GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCCCTGATC 377  
QY 361 CAGCAATGCCGCGGTGAGTAAGAGGCTTTCGGGTTGTAAGCTCTTTTCAGTCGAGAGA 420  
DB 378 CAGCAATGCCGCGGTGAGTAAGAGGCTTTCGGGTTGTAAGCTCTTTTCAGTCGAGAGA 437  
QY 421 AAAGGTTACGTTAATATCGTGAATCATGACGCTATCGACAGAGAGACACCGGCTTAC 480  
DB 438 AAAGGTTGAGTAATATATCACAACTTATGACGCTACCAAGAGAGACACCGGCTTAC 497  
QY 481 TACGCGCCAGCAGCGCGGCTTAATAGCGGTGCAAGCGTTATCGGAATTTCTGGGCGT 540  
DB 498 TACGCGCCAGCAGCGCGGCTTAATAGCGGTGCAAGCGTTATCGGAATTTCTGGGCGT 557  
QY 541 AAAGGTTGCCAGCGCGCTTTTGAATCGATGTAATCCCGCGGCTTAACTGGGGAAT 600  
DB 558 AAAGGTTGCCAGCGCGCTTTTGAATCGATGTAATCCCGCGGCTTAACTGGGGAAT 617  
QY 601 TCGGTTTGAATCAAGGCTGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660  
DB 618 TCGGTTTGAATCAAGGCTGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 677  
QY 661 TGAATGCGTGAAGATGAGAGAAATCATGATGCGAAGCGCCCTCCGGGTTAACT 720  
DB 678 TGAATGCGTGAAGATGAGAGAAATCATGATGCGAAGCGCCCTCCGGGTTAACT 737  
QY 721 GACGCTCATGACGAAGCGTGGGAGCAACAGATTAACCTGTGTAGTCCACGCC 780  
DB 738 GACGCTCATGACGAAGCGTGGGAGCAACAGATTAACCTGTGTAGTCCACGCC 797

QY 781 CTAACGATGTCATTAATGTTGGGCGCTTATTAAGGCTTGGTAACGAAGCTAAGCGCTGA 840  
DB 798 CTAACGATGTCATTAATGTTGGGCGCTTATTAAGGCTTGGTAACGAAGCTAAGCGCTGA 857  
QY 841 AGTTGACCGCGCTGGGAGTACGCTGCGCAAGATTAACTCAAGAGATTGACGGGAGCC 900  
DB 858 AGTTGACCGCGCTGGGAGTACGCTGCGCAAGATTAACTCAAGAGATTGACGGGAGCC 917  
QY 901 GCACAGCGGTGATTAATGTGATTAATTCATGATGCAACGCGAATACTTACCTT 960  
DB 918 GCACAGCGGTGATTAATGTGATTAATTCATGATGCAACGCGAATACTTACCTT 977  
QY 961 GACATGTGCGCAATTTTCTAGATAGATTAGTG-CTTCGGGAAACGCTTACACAGGTGCT 1019  
DB 978 GACATGTGCGCAATTTTCTAGATAGATTAGTG-CTTCGGGAAACGCTTACACAGGTGCT 1037  
QY 1020 GCATGCTGTGCTGACGCTGCTGCTGAGATGTTGGTTAAGTCCCGCAAGCGCAAC 1079  
DB 1038 GCATGCTGTGCTGACGCTGCTGCTGAGATGTTGGTTAAGTCCCGCAAGCGCAAC 1097  
QY 1080 CTTGTCTAATTAATGCGCATATTGTTGGGCACTTTAATGACATGCGCGTGACAAAC 1139  
DB 1098 CTTGTCTAATTAATGCGCATATTGTTGGGCACTTTAATGACATGCGCGTGACAAAC 1157  
QY 1140 GGAAGAGGTGGGATGACGTCAGTCTCATGCGCCCTTAAGGGTAAAGGCTTCAACAGTA 1199  
DB 1158 GGAAGAGGTGGGATGACGTCAGTCTCATGCGCCCTTAAGGGTAAAGGCTTCAACAGTA 1217  
QY 1200 ATACATGCGCGCTTACAGAGGTTGCCAACC CGCAGGAGGAGCTAATCTCAGAAAGCGC 1259  
DB 1218 ATACATGCGCGCTTACAGAGGTTGCCAACC CGCAGGAGGAGGAGCTAATCTCAGAAAGCGC 1277  
QY 1260 GTCTGATGTCGAGATCGAGTCTGCAACTCGGTGAATCGGAATTCCTAGTAATCG 1319  
DB 1278 GTCTGATGTCGAGATCGAGTCTGCAACTCGGTGAATCGGAATTCCTAGTAATCG 1337  
QY 1320 CGGATCAGCATGTGCGGCTGAAATCGTTCGGGGCTTGTACACACGCGCGCTCACCA 1379  
DB 1338 CGGATCAGCATGTGCGGCTGAAATCGTTCGGGGCTTGTACACACGCGCGCTCACCA 1397  
QY 1380 TGGAGTGGGTTTCCAGGAAGCAGTGTCTAAACCGTAAGAGAGCGCTTTCACAGGTG 1439  
DB 1398 TGGAGTGGGTTTCCAGGAAGCAGTGTCTAAACCGTAAGAGAGCGCTTTCACAGGTG 1457  
QY 1440 AGATTCAATGACTGGGCTG 1457  
DB 1458 AGATTCAATGACTGGGCTG 1475

RESULT 13  
AJ621032  
LOCUS AJ621032 1445 bp DNA linear BCT 08-JUN-2004  
DEFINITION Nitrosomonas sp. Is943 16S rRNA gene, isolate Is943.  
ACCESSION AJ621032.1 GI:40994851  
VERSION 16S ribosomal RNA; 16S rRNA gene.  
KEYWORDS Nitrosomonas sp. Is943  
SOURCE Nitrosomonas sp. Is943  
ORGANISM Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1  
AUTHORS Bollmann, A., Schmidt, I. and Bodelier, P. L. E.  
TITLE Influence of salt on activity, growth, protein pattern and fatty acid composition of the new isolated ammonia oxidizing strain Nitrosomonas strain Is943  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1445)  
AUTHORS Bollmann, A.  
TITLE Direct Submission  
JOURNAL Submitted (15-JUN-2004) Bollmann A., Department for Microbial Ecology, NIO-KNMI Centre for Limnology, Rijksstraatweg 6, 3631 AC Nieuwerlands, NETHERLANDS

FEATURES  
source  
Location/Qualifiers  
1. .1445  
/organism="Nitrosomonas sp. Is343"  
/mol\_type="genomic DNA"  
/isolate="Is343"  
/isolation\_source="Brazilian water"  
/db\_xref="taxon:261298"  
/country="Netherlands:River Schelde"  
1. .144  
/gene="16S rRNA"  
1. .144  
/gene="16S rRNA"  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 95.4%; Score 1390.6; DB 1; Length 1445;  
Best Local Similarity 98.5%; Pred. No. 9.1e-15;  
Matches 1424; Conservative 0; Mismatches 20; Indels 2; Gaps 2;  
QY 5 AACGCTGGCGGCATGCTTTACATGCAAGTCGACGCGACGCGATGCTTGCACTGG 64  
DB 1 AAGCGTGGGGGATGCTTTACATGCAAGTCGACGCGACGCGATGCTTGCACTGG 60  
QY 65 TGGCGAGTGGCGGACGGGTGATGATGCAATGCGAAGCTATCCGAAAGAGGGGGTTAACG 124  
DB 61 TGGCGAGTGGCGGACGGGTGATGATGCAATGCGAAGCTATCCGAAAGAGGGGGTTAACG 120  
QY 125 ATCGAAGATGCTTAATCCGATATACCTTAAGAGGAAACAGGGGATCGAAAGAC 184  
DB 121 ATCGAAGATGCTTAATCCGATATACCTTAAGAGGAAACAGGGGATCGAAAGAC 180  
QY 185 TTGCGCTTTTGGACGCGCCGATGCTGATGCTAGTTGGTGGGTAAAGGCTTACCAAG 244  
DB 181 TTGCGCTTTTGGACGCGCCGATGCTGATGCTAGTTGGTGGGTAAAGGCTTACCAAG 240  
QY 245 GCGACATCAATGTTGGTCTGAGAGGACGACGACGACCTGGGACTGAGACAGCGCC 304  
DB 241 GCGACATCAATGTTGGTCTGAGAGGACGACGACGACCTGGGACTGAGACAGCGCC 300  
QY 305 AGACTCTTACGGAGGACGAGTGGGGAATTTTGAATGGGCGCAAGGCTGATCCAG 364  
DB 301 AGACTCTTACGGAGGACGAGTGGGGAATTTTGAATGGGCGCAAGGCTGATCCAG 360  
QY 365 AATGCGCGTGAAGAGGCTTCGGGTTGAAAGCTCTTCACTGCAAGAAAG 424  
DB 361 AATGCGCGTGAAGAGGCTTCGGGTTGAAAGCTCTTCACTGCAAGAAAG 419  
QY 425 GTTACGGTAAATATCTGACTCATGACGGTATCGACAGAAAGACCGGCTTAACG 484  
DB 420 GTTACGGTAAATATCTGACTCATGACGGTATCGACAGAAAGACCGGCTTAACG 479  
QY 485 TGGCAGAGCGCGGTTAATCTGAGGGTGAAGGTTAATCGGAATTAAGGGGCTTAAG 544  
DB 480 TGGCAGAGCGCGGTTAATCTGAGGGTGAAGGTTAATCGGAATTAAGGGGCTTAAG 539  
QY 545 GGTGCGGACGCGCTTGTAAAGTCAAGTGAATCCCGGCTTAACTTGGGATTTGCG 604  
DB 540 GGTGCGGACGCGCTTGTAAAGTCAAGTGAATCCCGGCTTAACTTGGGATTTGCG 599  
QY 605 TTTGAAACTCAAGAGCTAGAGTGTGCAAGAGGAGTGAATTCATGTGTAGCAGTGA 664  
DB 600 TTTGAAACTCAAGAGCTAGAGTGTGCAAGAGGAGTGAATTCATGTGTAGCAGTGA 659  
QY 665 ATGGGTAGATATGAAAGAAACATCGATGGGAAAGGAGCTCTGGGTTAAAGCTAGC 724  
DB 660 ATGGGTAGATATGAAAGAAACATCGATGGGAAAGGAGCTCTGGGTTAAAGCTAGC 719  
QY 725 CTGATGACGAAACGTTGGGAGCAACAGAGTTAGTACCCTGTAGTCCACGCGCTTAA 784  
DB 720 CTGATGACGAAACGTTGGGAGCAACAGAGTTAGTACCCTGTAGTCCACGCGCTTAA 779  
QY 785 ACGATGTCACTAGTTGTTGGGCTTTATTAAGCTTTGTAACGAAGCTTAACGCTGAAGTT 844

DB 780 ACGATGTCACTAGTTGTTGGGCTTTATTAAGCTTATGAAGAACTTAACGCTGAAGTT 839  
QY 845 GACCGCTGGGAGTACGCTGCGAAGATTTAAACCTCAAGAAATTGACGGGACCGGAC 904  
DB 840 GACCGCTGGGAGTACGCTGCGAAGATTTAAACCTCAAGAAATTGACGGGACCGGAC 899  
QY 905 AAGCGGTGATATGATGATTAATGATGACGACGCAAAACCTTACTACCTTTGACA 964  
DB 900 AAGCGGTGATATGATGATTAATGATGACGACGCAAAACCTTACTACCTTTGACA 959  
QY 965 TGTACGAAATTTTCTAGAGATGATTAAGTGTG-CTTGGGAAACGTTAACAGGCTGCAT 1023  
DB 960 TGTACGAAATTTTCTAGAGATGATTAAGTGTG-CTTGGGAAACGTTAACAGGCTGCAT 1019  
QY 1024 GCGTGTGCTGACCTGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTT 1083  
DB 1020 GCGTGTGCTGACCTGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTT 1079  
QY 1084 GTCATTAATGTCATCATTTGGTGGGCACTTAATGAGACTGCGGTGACAAACCGAG 1143  
DB 1080 GTCATTAATGTCATCATTTAGTTGGGCACTTAATGAGACTGCGGTGACAAACCGAG 1139  
QY 1144 GAAGTGGGAGTACGTCATGCTCATGATGCTTATGGGTAGGGCTTACACGTAATAC 1203  
DB 1140 GAAGTGGGAGTACGTCATGCTCATGATGCTTATGGGTAGGGCTTACACGTAATAC 1199  
QY 1204 AATGGCGCTACAGAGGTTGGCCAAACCGCGAGGGGAGCTAATCTCAAGAAAGCGGTG 1263  
DB 1200 AATGGCGCTACAGAGGTTGGCCAAACCGCGAGGGGAGCTAATCTCAAGAAAGCGGTG 1259  
QY 1264 TAGTCCGATTCGAGATCTGCACTGACCTCGTGAAGTGGAAATGCTAGTATCGCGA 1323  
DB 1260 TAGTCCGATTCGAGATCTGCACTGACCTCGTGAAGTGGAAATGCTAGTATCGCGA 1319  
QY 1324 TCAGCATGTCGCGGTGAATACGTTCCGGGTCTTGTACACACGCGCGTCAACATGAG 1383  
DB 1320 TCAGCATGTCGCGGTGAATACGTTCCGGGTCTTGTACACACGCGCGTCAACATGAG 1379  
QY 1384 AGTGGTTTACACAGAGAGTGTCTAATCCGTAAAGAGGCGCTTGGCAGGTGAAGT 1443  
DB 1380 AGTGGTTTACACAGAGAGTGTCTAATCCGTAAAGAGGCGCTTGGCAGGTGAAGT 1439  
QY 1444 TCATGA 1449  
DB 1440 TCATGA 1445  
RESULT 14  
AF272420 1450 bp DNA linear BCT 05-DEC-2000  
LOCUS Nitrosomonas aestuarii 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF272420  
ACCESSION AF272420  
VERSION AF272420.1 GI:11545280  
KEYWORDS  
SOURCE Nitrosomonas aestuarii  
ORGANISM Nitrosomonas aestuarii  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1 (bases 1 to 1450)  
AUTHORS Burkhold,U., Pommerening-Roseer,A., Juretschko,S., Schmid,M.C.,  
Koops,H.P. and Wagner,M.  
Phylogeny of all recognized species of ammonia oxidizers based on  
comparative 16S rRNA and amoA sequence analysis: implications for  
molecular diversity surveys  
APPL. Environ. Microbiol. 66 (12), 5368-5382 (2000)  
JOURNAL 2 (bases 1 to 1450)  
PUBMED Burkhold,U., Pommerening-Roseer,A., Juretschko,S., Schmid,M.C.,  
Koops,H.P. and Wagner,M.  
REFERENCES Direct Submission  
TITLE Submitted (26-MAY-2000) Lehrstuhl fuer Mikrobiologie, Technische  
JOURNAL Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES Location/Qualifiers



```
source
1.1450
/organism="Nitrosomonas aestuarii"
/mol_type="genomic DNA"
/db_xref="taxon:52441"
<1..>1450
/product="16S ribosomal RNA"

ORIGIN

Query Match 95.0%; Score 1383.6; DB 1; Length 1450;
Best Local Similarity 97.6%; Pred. No. 1.2e-14;
Matches 1415; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 2 TTGAACGCTGGGCGCATGCTTTACACATGCAAGTGCAGACGCGACGCGATGCTTGCATC 61
DB 1 TTGAACGCTGGGCGCATGCTTTACACATGCAAGTGCAGACGCGACGCGATGCTTGCATC 60
QY 62 TGGTGCGAGTGGCGGACGCGGTGAGTAATGATCCGAAAGTTCGAGAAAGAGGGGGTGA 121
DB 61 TGGTGCGAGTGGCGGACGCGGTGAGTAATGATCCGAAAGTTCGAGAAAGTGGGGGATGA 120
QY 122 CGCATCGAAAGATGCTCTAATACCGCATATCTCTAAGAGAAAGACGAGGATCGAAG 181
DB 121 CGCATCGAAAGATGCTCTAATACCGCATATCTCTAAGAGAAAGACGAGGATCGAAG 180
QY 182 ACCTTGGCGCTTTGGAGCGCGCATGCTGATTTAGCTAGTTGGGGGTAAAGGCGTACC 241
DB 181 ACCTTGGCGCTTTGGAGCGCGCATGCTGATTTAGCTAGTTGGGGGTAAAGGCGTACC 240
QY 242 AAGGCGACGATCAGTATGTTGCTGAGAGAGACGACGACACTGGAGCTGAGACAGG 301
DB 241 AAGGCGACGATCAGTATGTTGCTGAGAGAGACGACGACACTGGAGCTGAGACAGG 300
QY 302 CCCAAGCTCTTACGGGAGGACAGTGGGGAAATTTTGGACAATGGGCGCAACCTGATCC 361
DB 301 CCCAAGCTCTTACGGGAGGACAGTGGGGAAATTTTGGACAATGGGCGCAACCTGATCC 360
QY 362 AGCAATGCCGCGGTGAGTGAAGAGCGCTTGGGTTGTAAAGCTTTTCAGTGAAGAA 421
DB 361 AGCAATGCCGCGGTGAGTGAAGAGCGCTTGGGTTGTAAAGCTTTTCAGTGAAGAA 420
QY 422 AAGGTACCGGTAAATATATGATCATGACGCTATCGACAGAGAAAGACCGGCTAACT 481
DB 421 AAGGTACCGGTAAATATATGATCATGACGCTATCGACAGAGAAAGACCGGCTAACT 480
QY 482 ACGTCCAGACGACCGCGGTAAATCGTAGGGTGCAGAGCTTAAATCGGAATTAATCGGCGCTA 541
DB 481 ACGTCCAGACGACCGCGGTAAATCGTAGGGTGCAGAGCTTAAATCGGAATTAATCGGCGCTA 540
QY 542 AAGGATGGCGACGCGCTTTGTAAGTCAGATGTAAGAAATCCCGGGCTTAACTCGGGAAT 601
DB 541 AAGGATGGCGACGCGCTTTGTAAGTCAGATGTAAGAAATCCCGGGCTTAACTCGGGAAT 600
QY 602 GCGTTGAAATCTAACAAGGCTAGAGTGTGCAGAGGAGGTGGAATTCATGTGTAGCACT 661
DB 601 GCGTTGAAATCTAACAAGGCTAGAGTGTGCAGAGGAGGTGGAATTCATGTGTAGCACT 660
QY 662 GAAATGCGTAGATATGGAAGAACATCGATGCGGAGGACGCTCTCGGTTAAACACTG 721
DB 661 GAAATGCGTAGATATGGAAGAACATCGATGCGGAGGACGCTCTCGGTTAAACACTG 720
QY 722 ACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGTACCTCGTAGTCCAGCGCC 781
DB 721 ACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGTACCTCGTAGTCCAGCGCC 780
QY 782 TAAACGATGTCAATAGTGTGGGCTTTATTAAGCTTGTGTAACGAAGCTTAACGCGTGA 841
DB 781 TAAACGATGTCAATAGTGTGGGCTTTATTAAGCTTGTGTAACGAAGCTTAACGCGTGA 840
QY 842 GTTGAACCGCTGGGAGTACGCTCGCAAGATTAAACCTAAAGAAATTGAAGGAGACCGC 901
DB 841 GTTGAACCGCTGGGAGTACGCTCGCAAGATTAAACCTAAAGAAATTGAAGGAGACCGC 900
QY 902 CACAAGCGGTGATTATGTGATTAATTGATGCAACGGAAGAAACCTTAACCTACCTTG 961
```

```
DB 901 CACAAGCGGTGATTATGTGATTAATTGATGCAACGGAAGAAACCTTAACCTACCTTG 960
QY 962 ACATGTACGGAATTTTCTAGAGATGATTAGTG-CTTCGGGAAGCTTAACACAGTGTG 1020
DB 961 ACATGTACGGAATTTTCTAGAGATGATTAGTG-CTTCGGGAAGCTTAACACAGTGTG 1020
QY 1021 CATGGCTGTGCTGACGCTGTGTGAGATGATTGGGTAAATCCCGCAAGAGCGCAAC 1080
DB 1021 CATGGCTGTGCTGACGCTGTGTGAGATGATTGGGTAAATCCCGCAAGAGCGCAAC 1080
QY 1081 CTGTCTAATTAATGCGCATCATTTGGTGGGCACTTAAATGAGACTGCCGTGACAAACCG 1140
DB 1081 CTGTCTAATTAATGCGCATCATTTGGTGGGCACTTAAATGAGACTGCCGTGACAAACCG 1140
QY 1141 GAGGAAGGTGGGATGATGATGCTCAAGTCTCATAGGCCCTTATGGGTGAGGCTTACACGTA 1200
DB 1141 GAGGAAGGTGGGATGATGATGCTCAAGTCTCATAGGCCCTTATGGGTGAGGCTTACACGTA 1200
QY 1201 TACAATGGCGGTACAGAGGTTGCAACCGCGAGGGGAGCTTAATCTCAGAAAGCGCG 1260
DB 1201 TACAATGGCGGTACAGAGGTTGCAACCGCGAGGGGAGCTTAATCTCAGAAAGCGCG 1260
QY 1261 TCGTAGTCGGATCGGATCTGCAACTGCACTCGGTGAAGTCGGAATGCTAGTAATGCG 1320
DB 1261 TCGTAGTCGGATCGGATCTGCAACTGCACTCGGTGAAGTCGGAATGCTAGTAATGCG 1320
QY 1321 GGATCAGATATGCTCGGCTGAATAGCTTCCGGGCTTGTATACACACCGCCCTCACACAT 1380
DB 1321 GGATCAGATATGCTCGGCTGAATAGCTTCCGGGCTTGTATACACACCGCCCTCACACAT 1380
QY 1381 GGGAGTGGGTTTCAACAGACAGTATGCTTAACGTAAGAGAGGAGCGCTTCCACGCTGA 1440
DB 1381 GGGAGTGGGTTTCAACAGACAGTATGCTTAACGTAAGAGAGGAGCGCTTCCACGCTGA 1440
QY 1441 GATTCAATGAC 1450
DB 1441 GATTCAATGAC 1450

RESULT 15
AF272424 1513 bp DNA linear BCT 05-DEC-2000
LOCUS AF272424
DEFINITION Nitrosomonas sp. NM 51 16S ribosomal RNA gene, partial sequence.
ACCESSION AF272424
VERSION AF272424.1 GI:11545284
KEYWORDS
SOURCE
ORGANISM Nitrosomonas sp. Nm51
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosomonas.
REFERENCE
1 (bases 1 to 1513)
Pukhold, U., Pommerening-Rosser, A., Juretschko, S., Schmid, M.C.,
Koops, H.P. and Wagner, M.
Phylogeny of all recognized species of ammonia oxidizers based on
comparative 16S rRNA and amoA sequence analysis: implications for
molecular diversity surveys
JOURNAL Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)
PUBMED 11097916
REFERENCE
2 (bases 1 to 1513)
Pukhold, U., Pommerening-Rosser, A., Juretschko, S., Schmid, M.C.,
Koops, H.P. and Wagner, M.
Direct Submission
JOURNAL Submitted (26-MAY-2000) Lehrstuhl fuer Mikrobiologie, Technische
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES
source
1.1513
/organism="Nitrosomonas sp. Nm51"
/mol_type="genomic DNA"
/db_xref="taxon:133720"
<1..>1513
/product="16S ribosomal RNA"

ORIGIN
```





Db 1 AGCTGGGGGCGGCTTTACACATGCAAGTCGAAAGGCGAGCAAGGGTCTTGACCTGGT 60  
Qy 66 GCGGATGGCGGACGGGTGATGTAATGCATCGGAACGTATCCGAAGAGGGGGGTAAACGA 125  
Db 61 GCGGATGGCGGACGGGTGATGTAATGCATCGGAACGTATCCGAAGAGGGGGGTAAACGA 120  
Qy 126 TCGAAGATGCTGTAATACCGATATATCTTAAGGAGGAAAGCAGGGGATCGAAAGACT 185  
Db 121 TCGAAGATGCTGTAATACCGATATATCTTAAGGAGGAAAGCAGGGGATCGAAAGACT 180  
Qy 186 TCGGCTTTTGGAGCGGCGCGATGTCGATTAAGTGTGGGTAAAGGCTTACCAAG 245  
Db 181 TCGGCTTTTGGAGTGGCGCGATGCTGATTAAGTGTGGGTAAAGGCTTACCAAG 240  
Qy 246 CGACGATCAGTATGTTGCTTGAGAGGACGACCGGACCTGGAGTGAACAAGGCCA 305  
Db 241 CGACGATCAGTATGTTGCTTGAGAGGACGACCGGACCTGGAGTGAACAAGGCCA 300  
Qy 306 GACTCTACGGGAGGAGGAGGAGGAGTTTGAACAATGGGCGGAAGCTGATCGAGA 365  
Db 301 GACTCTACGGGAGGAGGAGGAGGAGTTTGAACAATGGGCGGAAGCTGATCGAGA 360  
Qy 366 ATGCGCGGTGATGTAAGAGGCTTCGGGTGTAAAGCTCTTCACTGAGAGGAAAG 425  
Db 361 ATGCGCGGTGATGTAAGAGGCTTCGGGTGTAAAGCTCTTCACTGAGAGGAAAG 420  
Qy 426 TTAAGGTAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 485  
Db 421 TTAAGGTAATATATCAATATGATGATGATGATGATGATGATGATGATGATGAT 480  
Qy 486 GCGACAGCGCGGCTTAATGATGAGGTCGAAGGCTTAATCGGAATTAAGGCGTAAAG 545  
Db 481 GCGACAGCGCGGCTTAATGATGAGGTCGAAGGCTTAATCGGAATTAAGGCGTAAAG 540  
Qy 546 GTGCGCAGCGGCTTGTGAAGTCAATGTGAATCCCGGCGCTTAAGCTGGAATTCGT 605  
Db 541 GTGCGCAGCGGCTTGTGAAGTCAATGTGAATCCCGGCGCTTAAGCTGGAATTCGT 600  
Qy 606 TTGAAATCTAAGGCTTAAGGTCGAGAGGAGTGAATTCATGTGTACATGTAAG 665  
Db 601 TTGAAATCTAAGGCTTAAGGTCGAGAGGAGTGAATTCATGTGTACATGTAAG 660  
Qy 726 TCATCAGCAAGAAAGCGTGGGAGCAACAGGATTAAGTACCTGTGATGTCACGCTTAA 785  
Db 721 TCATCAGCAAGAAAGCGTGGGAGCAACAGGATTAAGTACCTGTGATGTCACGCTTAA 780  
Qy 786 CGATGTCATGTTGTTGGGCTTAATGAGCTTGGTGAAGGCTTAAGGCTTGAAGTTG 845  
Db 781 CGATGTCATGTTGTTGGGCTTAATGAGCTTGGTGAAGGCTTAAGGCTTGAAGTTG 840  
Qy 846 ACCGCTGGGGAGTACGGTGCAGAGTTAAACTCAAGAGATTTGACGGGGAGCCGAC 905  
Db 841 ACCGCTGGGGAGTACGGTGCAGAGTTAAACTCAAGAGATTTGACGGGGAGCCGAC 900  
Qy 906 AGCGGTGATTAATGATTAATTCGATGCAACGGGAAACCTTAACCTACCTTGAAT 965  
Db 901 AGCGGTGATTAATGATTAATTCGATGCAACGGGAAACCTTAACCTACCTTGAAT 960  
Qy 966 GTAGGAAATTTCTGAGATTAATTAAGTCTTGGGAAAGCTTAACAGGCTGCTGATG 1024  
Db 961 GTAGGAAATTTCTGAGATTAATTAAGTCTTGGGAAAGCTTAACAGGCTGCTGATG 1020  
Qy 1025 GCTGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1084  
Db 1021 GCTGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
Qy 1085 TCATTAATTTGCAATTTGTTGGGCACTTTAATGAGATGCGGCTGCAAAACCGGAG 1144  
Db 1081 TCATTAATTTGCAATTTGTTGGGCACTTTAATGAGATGCGGCTGCAAAACCGGAG 1140

Qy 1145 AAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1204  
Db 1141 AAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
Qy 1205 ATGCGCGCTGACAGAGGTTTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCTGT 1264  
Db 1201 ATGCGCGCTGACAGAGGTTTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCTGT 1260  
Qy 1265 AGTCCGATCGGAGTCTGCAACTGATCTCGTGAAGTGGAAATGCTAATGCTGCGAT 1324  
Db 1261 AGTCCGATCGGAGTCTGCAACTGATCTCGTGAAGTGGAAATGCTAATGCTGCGAT 1320  
Qy 1325 CAGATGTCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1384  
Db 1321 CAGATGTCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
Qy 1385 GTGGTTTCAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1444  
Db 1381 GTGGTTTCAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440  
Qy 1445 CATGACTGGGCTG 1457  
Db 1441 CATGACTGGGCTG 1453

RESULT 17  
CQ796926 1494 bp DNA linear PAT 19-APR-2004  
LOCUS  
DEFINITION Sequence 19 from Patent WO2004026772.  
ACCESSION CQ796926  
VERSION CQ796926.1 GI:46408552  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. 1494  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="N. Aestuarii-like AOB p4clone31 16S rDNA"

ORIGIN  
Query Match 93.3%; Score 1359.2; DB 6; Length 1494;  
Best Local Similarity 96.3%; Pred. No. 2.7e-14;  
Matches 1402; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

Qy 1 ATGGAAGCTGCGGATGCTTACACATGCAAGTCGAAAGGCGAGCAAGGCTTGCAT 60  
Db 19 ATTGAAGCTGCGGATGCTTACACATGCAAGTCGAAAGGCGAGCAAGGCTTGCAT 78  
Qy 61 CTGGTGGAGTGGCGGAGCGGCTGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 79 CTGGTGGAGTGGCGGAGCGGCTGATGATGATGATGATGATGATGATGATGATGAT 138  
Qy 121 ACGCATGGAAGATGCTAATACCGCATTAATCTTAAGAGGAAAGCAAGGAGTTCGAA 180  
Db 139 ACGCATGGAAGATGCTAATACCGCATTAATCTTAAGAGGAAAGCAAGGAGTTCGAA 198  
Qy 181 GACCTTGCGCTTTTGGAGCGGCGATGCTGATTAAGCTGATGATGATGATGATGATG 240  
Db 199 GACCTTGCGCTTTTGGAGCGGCGATGCTGATTAAGCTGATGATGATGATGATGATG 258  
Qy 241 CAAAGCGATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
Db 259 CAAAGCGATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 318

QY 301 GCCCAGACTCTTACCGGAGCAGCAGTGGGAAATTTTGACAAATGGGCGCAAGCCTGATC 360  
 DB 319 GCCCAGACTCTTACCGGAGCAGCAGTGGGAAATTTTGACAAATGGGCGCAAGCCTGATC 378  
 QY 361 CAGCAATGCCGGTGAATGAGAAAGCCCTTCCGGTTGTAAAGTCTTTTCACTGAGAAAGA 420  
 DB 379 CAGCAATGCCGGTGAATGAGAAAGCCCTTCCGGTTGTAAAGTCTTTTCACTGAGAAAGA 438  
 QY 421 AAAAGTTAGGTAATATCTGACTCATGACGGTATCGACAGAAAGACCGGCTAAC 480  
 DB 439 AAAAGTTAGTAAATATCAACTTATGACGGTATCGACAGAAAGACCGGCTAAC 498  
 QY 481 TAGTCCAGCAGCCCGGTAAATACGAGGTGCAAGCCTTATCGAAATTAATCTGGGCGT 540  
 DB 499 TAGTCCAGCAGCCCGGTAAATACGAGGTGCAAGCCTTATCGAAATTAATCTGGGCGT 558  
 QY 541 AAAAGGTGCGCAGCGGCTTTGTAAGTCAATGTGAATCCCGGGCTTAACTCTGGGAAT 600  
 DB 559 AAAAGGTGCGCAGCGGCTTTGTAAGTCAATGTGAATCCCGGGCTTAACTCTGGGAAT 618  
 QY 601 TCGGTTTGAATCAACAAGCTAGAGTGGCAGAGGAGGTGAATTCATGTGTAGCAG 660  
 DB 619 TCGGTTTGAATCAACAAGCTAGAGTGGCAGAGGAGGTGAATTCATGTGTAGCAG 678  
 QY 661 TGAATGCTAGATATGGAAGACATCGATGCGAAGCAGCCTCTGGGTTAACT 720  
 DB 679 TGAATGCTAGATATGGAAGACATCGATGCGAAGCAGCCTCTGGGTTAACT 728  
 QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGTATCCAGCC 780  
 DB 739 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGTATCCAGCC 798  
 QY 781 CTAAAGATATCACTAGTTGTTGGGCTTATTAAGCTTGGTAAAGCAAGCTAACCGGTGA 840  
 DB 799 CTAAAGATATCACTAGTTGTTGGGCTTATTAAGCTTGGTAAAGCAAGCTAACCGGTGA 858  
 QY 841 AGTTGACCGCTGGGAGTACCGTCCGCAAGATTAAACTCAAGAAATTTGACGGGAGCC 900  
 DB 859 AGTTGACCGCTGGGAGTACCGTCCGCAAGATTAAACTCAAGAAATTTGACGGGAGCC 918  
 QY 901 GCAACAGCGGTGATTAATGTGATTAATTTGATGCAACCGGAAACCTTAACCTTACCTT 960  
 DB 919 GCAACAGCGGTGATTAATGTGATTAATTTGATGCAACCGGAAACCTTAACCTTACCTT 978  
 QY 961 GACATGAGGAATTTTCTAGAGTATGATTAAGT-CTTGGGAAACGCTTAACAGAGTGT 1019  
 DB 979 GACATGAGGAATTTTCTAGAGTATGATTAAGT-CTTGGGAAACGCTTAACAGAGTGT 1038  
 QY 1020 GCAATGCTGTCTGACGCTGTCTGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGGAC 1079  
 DB 1039 GCAATGCTGTCTGACGCTGTCTGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGGAC 1098  
 QY 1080 CCTTGTCAATTAATTTGCAATTTTGTGGGCACTTTAATGAGACTCCCGGTGCAAAAC 1139  
 DB 1099 CCTTGTCAATTAATTTGCAATTTTGTGGGCACTTTAATGAGACTCCCGGTGCAAAAC 1158  
 QY 1140 GAGAGAAAGTGGGAGTACGTAAGTCTCAATGAGCCCTTATGAGGTAGGGCTTACAGATA 1199  
 DB 1159 GAGAGAAAGTGGGAGTACGTAAGTCTCAATGAGCCCTTATGAGGTAGGGCTTACAGATA 1218  
 QY 1200 ATACCAATGGCGCGTACAGAGGTTGCCAACCCCGGAGAGGAGCTAATCTCAGAAACGCG 1259  
 DB 1219 ATACCAATGGCGCGTACAGAGGTTGCCAACCCCGGAGAGGAGCTAATCTCAGAAACGCG 1278  
 QY 1260 GTCTAGTCCGAGTCGAGTCTGCAACTCGACTCGTGAAGTCCGAAATCGCTAATAATCG 1319  
 DB 1279 GTCTAGTCCGAGTCGAGTCTGCAACTCGACTCGTGAAGTCCGAAATCGCTAATAATCG 1338  
 QY 1320 CGGATTCAGCAATGTCGGGTGAATACGTTCCGGGCTTTTGTACACACCGCCGTCACACCA 1379  
 DB 1339 CGGATTCAGCAATGTCGGGTGAATACGTTCCGGGCTTTTGTACACACCGCCGTCACACCA 1398

QY 1380 TGGGAGTGGGTTTCAACGAAGCAGTAGTCTTAACCGTAAGGAGGCGCTTCCACGGTG 1439  
 DB 1399 TGGAGTTGGCTGCAACGAAAGTATGTTGTCTAACCTTCGGAGAGACGTTTACACGGTG 1458  
 QY 1440 AGATTCAATGACTGGGG 1455  
 DB 1459 TGGTCAATGACTGGGG 1474  
 RESULT 18  
 CQ796925 1467 bp DNA linear PAT 19-APR-2004  
 LOCUS Sequence 18 from Patent WO2004026772.  
 DEFINITION  
 ACCESSION CQ796925  
 VERSION CQ796925.1 GI:46408551  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 REFERENCE 1. Hovanec, T. A.  
 AUTHORS Ammonia-oxidizing bacteria and methods of using and detecting  
 TITLE chesame  
 JOURNAL Patent: WO 2004026772-A 18 01-APR-2004;  
 Aquaria Inc. (US)  
 FEATURES  
 source Location/Qualifiers  
 1..1467  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:332644"  
 /note="N. Astuaris-like AOB Paclone42 16S rDNA"  
 ORIGIN  
 Query Match 93.3%; Score 1358.8; DB 6; Length 1467;  
 Best Local Similarity 97.2%; Pred. No. 2.8e-14;  
 Matches 1414; Conservative 0; Mismatches 37; Indels 3; Gaps 3;  
 QY 1 ATTGAACGCTGGCGGATGCTTTTACATGACAGTGAACCGCAGACGGATGCTTGCAAT 60  
 DB 16 ATTGAACGCTGGCGGATGCTTTTACATGACAGTGAACCGCAGACGGATGCTTGCAAT 75  
 QY 61 CTGGTGGCAGATGGCGGAGCGGGTGAATGATCGGAACGTAATCGAAGAGGGGGGTA 120  
 DB 76 CTGGTGGCAGATGGCGGAGCGGGTGAATGATCGGAACGTAATCGAAGAGGGGGGTA 135  
 QY 121 ACCGATCGAAAGATGTCTAATACGCAATATCTTAAGAGGAAGACAGGGATGAAA 180  
 DB 136 ACCGATCGAAAGATGTCTAATACGCAATATCTTAAGAGGAAGACAGGGATGAAA 195  
 QY 181 GACCTTGGCTTTTGGAGCGGCCGATGTCTGATTAAGTGTGGGTAAAGGCTTAC 240  
 DB 196 GACCTTGGCTTTTGGAGCGGCCGATGTCTGATTAAGTGTGGGTAAAGGCTTAC 255  
 QY 241 CAAGGGAAGATAGTAAGTTGTCTGAGAGAGACAGCACTGGGACTGAGACAG 300  
 DB 256 CAAGGGAAGATAGTAAGTTGTCTGAGAGAGACAGCACTGGGACTGAGACAG 315  
 QY 301 GCCCAGACTCTTACCGGAGCAGCAGTGGGAAATTTTGACAAATGGGCGCAAGCCTGATC 360  
 DB 316 GCCCAGACTCTTACCGGAGCAGCAGTGGGAAATTTTGACAAATGGGCGCAAGCCTGATC 375  
 QY 361 CAGCAATGCCGGTGAATGAGAAAGCCCTTCCGGTTGTAAAGTCTTTTCACTGAGAAAGA 420  
 DB 376 CAGCAATGCCGGTGAATGAGAAAGCCCTTCCGGTTGTAAAGTCTTTTCACTGAGAAAGA 434  
 QY 421 AAAAGTTAGGTAATATCTGACTCATGACGGTATCGACAGAAAGACCGGCTAAC 480  
 DB 435 AAAAGTTAGTAAATATCAACTTATGACGGTATCGACAGAAAGACCGGCTAAC 494  
 QY 481 TAGTCCAGCAGCCCGGTAAATACGAGGTGCAAGCCTTATCGAAATTAATCTGGGCGT 540  
 DB 495 TAGTCCAGCAGCCCGGTAAATACGAGGTGCAAGCCTTATCGAAATTAATCTGGGCGT 554

```
QY 541 AAAGGTGGCAGCGCGCTTTGTAAGTCAAGTGTGAATCCCGGGCTTAACCTGGGAAT 600
|
|
|
Db 555 AAAGGTGGCAGCGCGCTTTGTAAGTCAAGTGTGAATCCCGGGCTTAACCTGGGAAT 614
|
|
|
QY 601 TCGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTAGCAG 660
|
|
|
Db 615 TCGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTAGCAG 674
|
|
|
QY 661 TGAATGCGTGAATATGGAAGAACATGATGGCCAGAGCGCTCTGGGTTAACT 720
|
|
|
Db 675 TGAATGCGTGAATATGGAAGAACATGATGGCCAGAGCGCTCTGGGTTAACT 734
|
|
|
QY 721 GAGCGTCAAGCAGAAAGGTGGGGAGCAACAGATTGATACCTGGTAGCCAGCGC 780
|
|
|
Db 735 GAGCGTCAAGCAGAAAGGTGGGGAGCAACAGATTGATACCTGGTAGCCAGCGC 794
|
|
|
QY 781 CTAAACGATGTCAACTAGTTGTGGGCTTAAATTAGGCTTGTAAACGACGCTGA 840
|
|
|
Db 795 CTAAACGATGTCAACTAGTTGTGGGCTTAAATTAGGCTTGTAAACGACGCTGA 854
|
|
|
QY 841 AGTTGACCGCTGGGGAGTACCGTCCGAAAGATTAAACCTAAAGATTGACGGGAGCC 900
|
|
|
Db 855 AGTTGACCGCTGGGGAGTACCGTCCGAAAGATTAAACCTAAAGATTGACGGGAGCC 914
|
|
|
QY 901 GACAAAGCGGTGGAATATGATGATTAATTCGATGCAACGCGAAACCTTACCTT 960
|
|
|
Db 915 GACAAAGCGGTGGAATATGATGATTAATTCGATGCAACGCGAAACCTTACCTT 974
|
|
|
QY 961 GACATGTACGCAATTTTCTAGAGATAGATTAGTC- TTCGGGAACGCTTAACAAGGTGCT 1019
|
|
|
Db 975 GACATGTACGCAATTTTCTAGAGATAGATTAGTC- TTCGGGAACGCTTAACAAGGTGCT 1034
|
|
|
QY 1020 GCATGGCTGCTCGTCACTGCTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGAAC 1079
|
|
|
Db 1035 GCATGGCTGCTCGTCACTGCTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGAAC 1094
|
|
|
QY 1080 CCTTGCAATTAATGTCATCATTTGGTGGGCACTTAATGAGACTGCGGTTGACAAAC 1139
|
|
|
Db 1095 CCTTGCAATTAATGTCATCATTTGGTGGGCACTTAATGAGACTGCGGTTGACAAAC 1154
|
|
|
QY 1140 GGAGGAAGGTGGGGAATGAGTCAGTCAAGTCTCATAGCCCTTAATGGGTAGGGCTTACACGTA 1199
|
|
|
Db 1155 GGAGGAAGGTGGGGAATGAGTCAGTCAAGTCTCATAGCCCTTAATGGGTAGGGCTTACACGTA 1214
|
|
|
QY 1200 ATACATAGCGCGTACAGAGGGTTCGCAACCCGCGAGGGGAGCTAATTCAGAAAGCGC 1259
|
|
|
Db 1215 ATACATAGCGCGTACAGAGGGTTCGCAACCCGCGAGGGGAGCTAATTCAGAAAGCGC 1274
|
|
|
QY 1260 GTGTAAGTCCGGAATGGAATCTGCAACTGCACTCGTGAAGTCCGGAATGCTAGTAATG 1319
|
|
|
Db 1275 GTGTAAGTCCGGAATGGAATCTGCAACTGCACTCGTGAAGTCCGGAATGCTAGTAATG 1334
|
|
|
QY 1320 CGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTTGTGTAACAACCGCGCTCAACCA 1379
|
|
|
Db 1335 CGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTTGTGTAACAACCGCGCTCAACCA 1394
|
|
|
QY 1380 TGGGAGTGGGTTTACACAGAGAGTGTCTTAACCTGAAGAGGGCGCTTGCACAGGTG 1439
|
|
|
Db 1395 TGGGAGTGGGTTTACACAGAGAGTGTCTTAACCTGAAGAGGGCGCTTGCACAGGTG 1453
|
|
|
QY 1440 AGATTCAATGACTGG 1453
|
|
|
Db 1454 AGATTCAATGACTGG 1467
|
|
|
RESULT 19
AB000700 1501 bp DNA linear BCT 06-MAR-2001
DEFINITION Nitrosomonas sp. JL21 gene for 16S rRNA, partial sequence.
ACCESSION AB000700
VERSION AB000700.1 GI:3107909
KEYWORDS Nitrosomonas sp. JL21
SOURCE
```

```
ORGANISM Nitrosomonas sp. JL21
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosomonas.
REFERENCE
1 (sites)
AUTHORS Suwa,Y., Sumino,T. and Noto,K.
TITLE Phylogenetic relationships of activated sludge isolates of ammonia
oxidizers with different sensitivities to ammonium sulfate
JOURNAL J. Gen. Appl. Microbiol. 43, 373-379 (1997)
REFERENCE
2 (bases 1 to 1501)
AUTHORS Suwa,Y.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-1997) Yuichi Suwa, National Institute for
Resources & Environment, Ecological Chemistry and Microbiology
Division, 16-3 Onogawa, Tsukuba, Ibaraki 305, Japan
(E-mail:suwa@nre.go.jp, Tel:+81-298-58-8318, Fax:+81-298-58-8309)
Location/Qualifiers
FEATURES
source
1..1501
/organism="Nitrosomonas sp. JL21"
/mol_type="genomic DNA"
/isolate="JL21"
/db_xref="taxon:153949"
<1..>1501
/product="16S ribosomal RNA"
ORIGIN
Query Match 92.6%; Score 1349.6; DB 1; Length 1501;
Best Local Similarity 95.8%; Pred. No.3.8e-14;
Matches 1398; Conservative 0; Mismatches 55; Indels 3; Gaps 1;
QY 1 ATGGAACGCTGGCGCATGCTTTACACATGCAAGTCGACGCGACGATGCTTGAT 60
|
|
|
Db 21 ATTGAACGCTGGCGCATGCTTTACACATGCAAGTCGACGCGACGATGCTTGAT 80
|
|
|
QY 61 CTGTGGCGAGTGGCGGACGGGTGATGATGATCGAACTATCCAGAAAGGGGGTA 120
|
|
|
Db 81 TTGTGGCGAGTGGCGGACGGGTGATGATGATCGAACTATCGAACTATCGGGGTA 140
|
|
|
QY 121 ACGCATGGAAGATGTGTAATACCGCATATCTCTAAGAGGAAGCAAGCGGATTCGAA 180
|
|
|
Db 141 ACGCATGGAAGATGTGTAATACCGCATATCTCTAAGAGGAAGCAAGCGGATTCGAA 200
|
|
|
QY 181 GACCTTGGCTTTTGAAGCGGCGCGATGCTGATTAAGTAGTGGTGAAGGCTTAC 240
|
|
|
Db 201 GACCTTGGCTTTTGAAGCGGCGCGATGCTGATTAAGTAGTGGTGAAGGCTTAC 260
|
|
|
QY 241 CAAGCGCATCACTAGTGTGTTGAGAGACGACCACTGGACTGAGACACG 300
|
|
|
Db 261 CAAGCGCATCACTAGTGTGTTGAGAGACGACCACTGGACTGAGACACG 320
|
|
|
QY 301 GCCCAAGTCTCTACGGGAGGAGCAGTGGGGAAATTTTGAACAATGGGCGCAAGCTGATC 360
|
|
|
Db 321 GCCCAAGTCTCTACGGGAGGAGCAGTGGGGAAATTTTGAACAATGGGCGCAAGCTGATC 380
|
|
|
QY 361 CAGCAATCCCGTGAAGTGAAGAGGCCCTTCGGGTTGAAAGCTTTCACTCGAAGA 420
|
|
|
Db 381 CAGCAATCCCGTGAAGTGAAGAGGCCCTTCGGGTTGAAAGCTTTCACTCGAAGA 440
|
|
|
QY 421 AAAGTTACGGTAATTAATCGTACTATGACCGTATGACAGAAAGACCGGCTAAC 480
|
|
|
Db 441 AAAGTTACGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 500
|
|
|
QY 481 TACGTCAGAGCAGCGCGGTAAATCGTAGGGTGAAGCGTTAATCGGAATTAAGGGGCT 540
|
|
|
Db 501 TACGTCAGAGCAGCGCGGTAAATCGTAGGGTGAAGCGTTAATCGGAATTAAGGGGCT 560
|
|
|
QY 541 AAAGGTGGCAGCGCGCTTTGTAAGTCAAGTGTGAATCCCGGGCTTAACTGGGAAT 600
|
|
|
Db 561 AAAGGTGGCAGCGCGCTTTGTAAGTCAAGTGTGAATCCCGGGCTTAACTGGGAAT 620
|
|
|
QY 601 TCGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTAGCAG 660
|
|
|
Db 621 TCGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTAGCAG 680
|
|
|
```

QY 661 TGAATGCTAGATATGAAAGAACATCGATGGCGAAGCGCTCTGGGTTAACT 720  
 Db 681 TGAATGCTAGATATGAAAGAACATCGATGGCGAAGCGCTCTGGGTTAACT 740  
 QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTGATACCTGGTAGTCCAGCGC 780  
 Db 741 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTGATACCTGGTAGTCCAGCGC 800  
 QY 781 CTAAACGATGTAACCTAGTTGTTGGGCTTAATTAGGCTTGTAACGAACCTAACCGGTGA 840  
 Db 801 CTAAACGATGTAACCTAGTTGTTGGGCTTAATTAGGCTTGTAACCTAACCGGTGA 860  
 QY 841 AGTTGACCGCTGGGGAGTAACGCTCGCAAGATTAAACTCAAGAAATTGACGGGAGCC 900  
 Db 861 AGTTGACCGCTGGGGAGTAACGCTCGCAAGATTAAACTCAAGAAATTGACGGGAGCC 920  
 QY 901 GCAACAGCGGTGATTAATGATGATTAACTGATGCAACCGCAAAACCTTAACCTACCTT 960  
 Db 921 GCAACAGCGGTGATTAATGATGATTAACTGATGCAACCGCAAAACCTTAACCTACCTT 980  
 QY 961 GACATGAGGAAATTTTCTAGATAGATTAGTGT---CTTGGGAAAGCTTAACAGAGT 1017  
 Db 981 GACATGAGGAAATTTTCTAGATAGATTAGTGT---CTTGGGAAAGCTTAACAGAGT 1040  
 QY 1018 CTGCAATGCTGTGCTGACGCTGCTGCTGATGATGTTGGTTAAAGTCCCGACAGAGCGCA 1077  
 Db 1041 CTGCAATGCTGTGCTGACGCTGCTGCTGATGATGTTGGTTAAAGTCCCGACAGAGCGCA 1100  
 QY 1078 ACCCTGTCTATTATTGCTCATCTTATGTTGGGCACTTTAATGAGACTGCGCGGTGACAA 1137  
 Db 1101 ACCCTGTCTATTATTGCTCATCTTATGTTGGGCACTTTAATGAGACTGCGCGGTGACAA 1160  
 QY 1138 CCGGAGAAAGGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTAATGGGATGGGCTTCAACG 1197  
 Db 1161 CCGGAGAAAGGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTAATGGGATGGGCTTCAACG 1220  
 QY 1198 TAATACATGCGCGCTGACAGAGGATTGCAACCCCGAGGGGAGCTTAATCTCAAGAAAGC 1257  
 Db 1221 TAATACATGCGCGCTGACAGAGGATTGCAACCCCGAGGGGAGCTTAATCTCAAGAAAGC 1280  
 QY 1258 GCGCTGATGCTCGGATCGGAGTCTGCAACTGCACTCCGTGAATGCGGAATCGCTAGTAAT 1317  
 Db 1281 GCGCTGATGCTCGGATCGGAGTCTGCAACTGCACTCCGTGAATGCGGAATCGCTAGTAAT 1340  
 QY 1318 CCGGATCAGCATGCTGCGGATGATAGCTTCCGGGCTTGTACACACCGCCGTCACAC 1377  
 Db 1341 CCGGATCAGCATGCTGCGGATGATAGCTTCCGGGCTTGTACACACCGCCGTCACAC 1400  
 QY 1378 CATGGAGTGGGTTTACAGAGAGCAGTAGTCTTAACCGTAAGAGAGGCGCTTGCCACGG 1437  
 Db 1401 CATGGAGTGGGTTTACAGAGAGCAGTAGTCTTAACCGTAAGAGAGGCGCTTGCCACGG 1460  
 QY 1438 TGAAGATTCAATGACTGGGCTG 1457  
 Db 1461 TGAAGATTCAATGACTGGGCTG 1480

## RESULT 20

AY123797 1499 bp DNA linear BCT 23-SBP-2003  
 LOCUS AY123797  
 DEFINITION Nitrosomonas sp. Nm84 16S ribosomal RNA gene, partial sequence.  
 ACCESSION AY123797  
 VERSION AY123797.1 GI:24474426  
 KEYWORDS Nitrosomonas sp. Nm84  
 SOURCE Nitrosomonas sp. Nm84  
 ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosomonas.  
 REFERENCE 1 (bases 1 to 1499)  
 AUTHORS Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roser, A. and Koops, H.-P.  
 TITLE 16S rRNA and amcA-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal

of a new lineage within the nitrosomonads  
 Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
 JOURNAL 13130037  
 PUBMED 2 (bases 1 to 1499)  
 REFERENCE Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roser, A. and Wagner, M.  
 AUTHORS Direct Submission  
 TITLE Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universität München, Am Hochanger 4, Freising 85350, Germany  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source 1..1499  
 /organism="Nitrosomonas sp. Nm84"  
 /mol\_type="genomic DNA"  
 /strain="Nm84"  
 /db\_xref="taxon:200124"  
 <1..>1499  
 /product="16S ribosomal RNA"  
 ORIGIN  
 Query Match 92.5%; Score 1347.8; DB 1; Length 1499;  
 Best Local Similarity 95.5%; Pred. No. 4,1e-14;  
 Matches 1395; Conservative 2; Mismatches 60; Indels 3; Gaps 1;  
 QY 1 ATTTGAAGCTGGGGGCGCATCTTTACATGCAAGTGAACGCGACGCGATGCTTGCAT 60  
 Db 3 ATTTGAAGCTGGGGGCGCATCTTTACATGCAAGTGAACGCGACGCGATGCTTGCAT 62  
 QY 61 CTGTGGCGAGTGGCGGAGCGGTTAGTAATGCAATGCAATGCAATGCAATGCAATGCAAT 120  
 Db 63 TNGGTGGCGAGTGGCGGAGCGGTTAGTAATGCAATGCAATGCAATGCAATGCAATGCAAT 122  
 QY 121 AGCGATGGAAGATGCTTAATACCGCATATCTTAAGAGGAAGCGGGGATGGA 180  
 Db 123 AGCGATGGAAGATGCTTAATACCGCATATCTTAAGAGGAAGCGGGGATGGA 182  
 QY 181 GACCTTGCGCTTTTGGAGCGCGCATGCTGATTTAGCTGATTTGGTGGGTTAAAGGCTTAC 240  
 Db 183 GACCTTGCGCTTTTGGAGCGCGCATGCTGATTTAGCTGATTTGGTGGGTTAAAGGCTTAC 242  
 QY 241 CAAGCGAGCATAGTAATGTTGTTCTGAGAGAGACGACGACCACTGGGACTGAGACAGC 300  
 Db 243 CAAGCGAGCATAGTAATGTTGTTCTGAGAGAGACGACGACCACTGGGACTGAGACAGC 302  
 QY 301 GCCCAACTCTTACGCGGAGCGACAGTGGGAAATTTTGGACATATGGGCGCAAGCTGATC 360  
 Db 303 GCCCAACTCTTACGCGGAGCGACAGTGGGAAATTTTGGACATATGGGCGCAAGCTGATC 362  
 QY 361 CAGCAATGCCGCGTGAAGAAAGGCTTGGGTTTAAAGCTCTTTCAGTGGAGAGA 420  
 Db 363 CAGCAATGCCGCGTGAAGAAAGGCTTGGGTTTAAAGCTCTTTCAGTGGAGAGA 422  
 QY 421 AAAGATTACCGTAAATATCTGATCATAGACGTAATCGACAGAAAGACCGGCTAAC 480  
 Db 423 AAAGATTACCGTAAATATCTGATCATAGACGTAATCGACAGAAAGACCGGCTAAC 482  
 QY 481 TACGTGCAGACAGCCCGGTTAATAGTGGGTGCAAGCGTTAATCGGAATTAATCTGGCGGT 540  
 Db 483 TACGTGCAGACAGCCCGGTTAATAGTGGGTGCAAGCGTTAATCGGAATTAATCTGGCGGT 542  
 QY 541 AAAGGGTGGCGAGCGGCTTGTAGTCAATGTAATATCCCGGCTTAATCTGGGAAT 600  
 Db 543 AAAGGGTGGCGAGCGGCTTGTAGTCAATGTAATATCCCGGCTTAATCTGGGAAT 602  
 QY 601 TGGCTTTGAAACTACAGAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTACAG 660  
 Db 603 TGGCTTTGAAACTACAGAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTACAG 662  
 QY 661 TGAATGCGTAGATATGGAAGACATGCAATGGCGAAGCAAGCTTCTGGGTTAACT 720  
 Db 663 TGAATGCGTAGATATGGAAGACATGCAATGGCGAAGCAAGCTTCTGGGTTAACT 722  
 QY 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGGATTAGATACCTGTAGTCCAGCGC 780

Db 723 GACGCTCATGACGAAAGCGTGGGAGCAAAAGAGATTAGATACCCTGGTAGTCCACGCC 782  
Qy 761 CTAACGATGTCATAGTTGTTGGCCCTTATTAGGCTTGGTATGACAGACTAACCGCTGA 840  
Db 763 CTAACGATGTCATAGTTGTTGGCCCTTATTAGGCTTGGTATGACAGACTAACCGCTGA 842  
Qy 841 AGTTACCGCCCTGGGAGATACCGTGCAGATTTAAACCAAGAAATTGACGGGAGACC 900  
Db 843 ATTGACCGCCCTGGGAGATACCGTGCAGATTTAAACCAAGAAATTGACGGGAGACC 902  
Qy 901 GCACAGCGCGTGGATTTATGATGATTAATTGATGCAACGCGAAGAACTTACCTTACCCTT 960  
Db 903 GCACAGCGCGTGGATTTATGATGATTAATTGATGCAACGCGAAGAACTTACCTTACCCTT 962  
Qy 961 GACATGATGCAATTTTCTAGATGATTAAGT---CTTGGGAAACGCTTACACAGGTG 1017  
Db 963 GACATGATGCAATTTTCTAGATGATTAAGT---CTTGGGAAACGCTTACACAGGTG 1022  
Qy 1018 CTGCAATGGCTGCTGACGCTCGTGTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1077  
Db 1023 CTGCAATGGCTGCTGACGCTCGTGTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1082  
Qy 1078 ACCCTTGTCAATTAATTCCTCATTTGTTGGGCACTTTAATGAGACTGCGGTGACAA 1137  
Db 1083 ACCCTTGTCAATTAATTCCTCATTTGTTGGGCACTTTAATGAGACTGCGGTGACAA 1142  
Qy 1138 CCGGAGGAAGGTGGGAGATACGTCATAGTCTTATGAGGCTTATGAGGCTTACACAG 1197  
Db 1143 CCGGAGGAAGGTGGGAGATACGTCATAGTCTTATGAGGCTTATGAGGCTTACACAG 1202  
Qy 1198 TAAATCAATGGCGCGTACAGAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1257  
Db 1203 TAAATCAATGGCGCGTACAGAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1262  
Qy 1258 GCGGTGATGTCGGAATCGGAGTCTGCAACTCGACTCGGTAATCGGAATCGCTAAT 1317  
Db 1263 GCGGTGATGTCGGAATCGGAGTCTGCAACTCGACTCGGTAATCGGAATCGCTAAT 1322  
Qy 1318 CCGGATCAAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTCACACCGCCCGTACAC 1377  
Db 1323 CCGGATCAAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTCACACCGCCCGTACAC 1382  
Qy 1378 CATGGAGTGGGTTTCAACGAGAGAGTCTAATCCGTAAGAGAGGCGCTTGCACAG 1437  
Db 1383 CATGGAGTGGGTTTCAACGAGAGAGTCTAATCCGTAAGAGAGGCGCTTGCACAG 1442  
Qy 1438 TGAGATTCACTGACTGGGGTG 1457  
Db 1443 TGAGATTCACTGACTGGGGTG 1462  
RESULT 21  
LOCUS AY123798 1499 bp DNA linear BCT 23-SEP-2003  
DEFINITION Nitrosomonas sp. Nm86 16S ribosomal RNA gene, partial sequence.  
ACCESSION AY123798  
VERSION AY123798.1 GI:24474427  
KEYWORDS  
SOURCE Nitrosomonas sp. Nm86  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
REFERENCE  
AUTHORS 1. Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roesser, A. and  
Koops, H. P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
JOURNAL  
PUBMED 13130037  
REFERENCE 2 (bases 1 to 1499)  
AUTHORS Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roesser, A. and  
Wagner, M.

TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1. 1499  
/organism="Nitrosomonas sp. Nm86"  
/mol\_type="genomic DNA"  
/strain="Nm86"  
/db\_xref="taxon:200125"  
<1..>1499  
rRNA  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 92.5%; Score 1347; DB 1; Length 1499;  
Best Local Similarity 95.6%; Pred. No. 4.2e-14;  
Matches 1396; Conservative 0; Mismatches 61; Indels 3; Gaps 1;  
Qy 1 ATTGAACGCTGGCGCATGCTTTACATGCAAGTCGAACCGGACGACGATGCTTGAT 60  
Db 3 ATTGAACGCTGGCGCATGCTTTACATGCAAGTCGAACCGGACGACGATGCTTGAT 62  
Qy 61 CTGGTGGAGTGGCGGACGGGTGATGATGATCGGAACCTATCCAGAAAGGGGGTA 120  
Db 63 CNGGTGGAGTGGCGGACGGGTGATGATGATCGGAACCTATCCAGAAAGGGGGTA 122  
Qy 121 ACGCATGGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAA 180  
Db 123 ACGCATGGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAA 182  
Qy 181 GACCTTGGCTTTTGAACCGCGCCGATGCTGATTAAGTATGTTGGTGGGTTAAAGCCTTAC 240  
Db 183 GACCTTATGCTTTTGAACCGCGCCGATGCTGATTAAGTATGTTGGTGGGTTAAAGCCTTAC 242  
Qy 241 CAAGCGCATGCTGATGTTGCTGAGAGAGACGACGACCACTGGGATGAGACAG 300  
Db 243 CAAGCGCATGCTGATGTTGCTGAGAGAGACGACGACCACTGGGATGAGACAG 302  
Qy 301 GCCCAGACTCCTACGGGAGGAGCAGGTGGGAAATTTTGAACAATGGGCGCAAGCCTGATC 360  
Db 303 GCCCAGACTCCTACGGGAGGAGCAGGTGGGAAATTTTGAACAATGGGCGCAAGCCTGATC 362  
Qy 361 CAGCAATGCGCGTGAAGTGAAGAAAGCCCTTGGGTTGAAAGCTCTTCACTGAGAA 420  
Db 363 CAGCAATGCGCGTGAAGTGAAGAAAGCCCTTGGGTTGAAAGCTCTTCACTGAGAA 422  
Qy 421 AAGGTTACGGTAAATATGCTGACTCATGACGTTATCGAAGAGAAAGCACCAGCTAAC 480  
Db 423 AAGGTTACGGTAAATATGCTGACTCATGACGTTATCGAAGAGAAAGCACCAGCTAAC 482  
Qy 481 TACGGCCAGAGCGCGGCTAATACGTAAGGTTGCAAGGCTTAATCGGAATTAAGTGGGCT 540  
Db 483 TACGGCCAGAGCGCGGCTAATACGTAAGGTTGCAAGGCTTAATCGGAATTAAGTGGGCT 542  
Qy 541 AAGGTTGCGGACGCGGCTTTGTAAGTCAATGTAATCCCGGCTTAACTGGGAAT 600  
Db 543 AAGGTTGCGGACGCGGCTTTGTAAGTCAATGTAATCCCGGCTTAACTGGGAAT 602  
Qy 601 TCGGTTGAAACTTACAGAGCTGAGTGTGCGAGAGGGAGTGTGAATTCATGTGTAGAG 660  
Db 603 TCGGTTGAAACTTACAGAGCTGAGTGTGCGAGAGGGAGTGTGAATTCATGTGTAGAG 662  
Qy 661 TGAATGTGTGAGATATGGAAGAAATCATGATGCGAAGGAGGCTCCTGGGTTAAACT 720  
Db 663 TGAATGTGTGAGATATGGAAGAAATCATGATGCGAAGGAGGCTCCTGGGTTAAACT 722  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGTACCTGGTATGTCACGCC 780  
Db 723 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGTACCTGGTATGTCACGCC 782  
Qy 781 CTAACGATGTCATAGTTGTTGGGCTTATTTAGGCTTGTGTAACGAAGCTAACCGCTGA 840  
Db 783 CTAACGATGTCATAGTTGTTGGGCTTATTTAGGCTTGTGTAACGAAGCTAACCGCTGA 842

Qy	841	AGTTGACCGCTTGGGAGAGTACGGTCGCAAGATTAAACTCAAGGAATTGAACGGGACCC	900
Db	843	AGTTGACCGCTTGGGAGAGTACGGTCGCAAGATTAAACTCAAGGAATTGAACGGGACCC	902
Qy	901	GCACAAAGGGGAGATTATGTGGATTAAATTGATGCAACGCGAAGAAACCTTACCTACCTT	960
Db	903	GCACAAAGGGGAGATTATGTGGATTAAATTGATGCAACGCGAAGAAACCTTACCTACCTT	962
Qy	961	GACATGTACCAATTTTCTTAGAGATGATTAAGTCT--TCGGGAACGCTTAACACAGGTG	1017
Db	963	GACATGTACCAAAAGTTTTCAGAGATGAAATGTGTCCTGGAAGAGAAATCTGAACACAGGTG	1022
Qy	1018	CTGCATGGCTGTCTGACCTCGTGTCTGTGAGATGTTGGGTTAACTCCCCCAAGAGGCA	1077
Db	1023	CTGCATGGCTGTCTGACCTCGTGTCTGTGAGATGTTGGGTTAACTCCCCCAAGAGGCA	1082
Qy	1078	ACCCCTTGCAATTAATGGCATCTTGGGTTGGGCACTTAAAGACATGCGCGGTGACAA	1137
Db	1083	ACCCCTTGCAATTAATGGCATCTTAACTTAAGTGGGCACTTAAATGAGACTGCGCGGTGATMA	1142
Qy	1138	CCGAGAGAAAGTGAGGAGATGACGTCAAGTCTCATAGCCCTTAATGGGTAAGGCTTACACG	1197
Db	1143	CCGAGAGAAAGTGAGGAGATGACGTCAAGTCTCATAGCCCTTAATGGGTAAGGCTTACACG	1202
Qy	1198	TAAATCAATGCGCGCTTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAAAGC	1257
Db	1203	TAAATCAATGCGCGCTTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAAAGC	1262
Qy	1258	GGCTGTGATCCGGATTCGGAGTCTGCCAATCTGACCTCGGTGAAGTCGGAATGCTAGTAT	1317
Db	1263	GGCTGTGATCCGGATTCGGAGTCTGCCAATCTGACCTCGGTGAAGTCGGAATGCTAGTAT	1322
Qy	1318	CGCGGATCAGCATGTGCGCGTGAAATACCTTCCCGGCTTTGTACACACCGCCGCTCACAC	1377
Db	1323	CGCGGATCAGCATGTGCGCGTGAAATACCTTCCCGGCTTTGTGTACACACCGCCGCTCACAC	1382
Qy	1378	CATGGAAGTGGGTTTACACAGAACGAGTATCTAACCTTAAAGAGGGCGCTTGTCCACAG	1437
Db	1383	CATGGAAGTGGGTTTACACAGAACGAGTATCTAACCTTAAAGAGGGCGCTTGTCCACAG	1442
Qy	1438	TGAGATTCAATGACTGGGGTG 1457	
Db	1443	TGAGATTCAATGACTGGGGTG 1462	
RESULT 22			
AB186828		1499 bp	DNA linear ENV 11-AUG-2005
LOCUS	Uncultured bacterium gene for 16S rRNA, partial sequence.		
DEFINITION	AB186828 clone: TSAQ22.		
ACCESSION	AB186828.1 GI:51100881		
VERSION	ENV.		
KEYWORDS	uncultured bacterium		
SOURCE	uncultured bacterium		
ORGANISM	Bacteria; environmental samples.		
REFERENCE	1 Yoshida, N., Takahashi, N. and Hiraiishi, A.		
AUTHORS	Phylogenetic characterization of a polychlorinated-dioxin-dechlorinating microbial community by use of microcosm studies		
TITLE	Appl. Environ. Microbiol. 71 (8), 4325-4334 (2005)		
JOURNAL	16085820		
PUBMED	2 (bases 1 to 1499)		
REFERENCE	Hiraiishi, A. and Yoshida, N.		
AUTHORS	Direct Submission		
TITLE	Submitted (06-AUG-2004) Akira Hiraiishi, Toyohashi University of		
JOURNAL	Technology, Department of Ecological Engineering, Tempaku-cho,		
	Toyohashi, Aichi, 441-8580, Japan (E-mail:hiraiishi@eco.tut.ac.jp,		
	Tel:81-532-44-6913, Fax:81-532-44-6929)		
FEATURES	Location/Qualifiers		
source	1..1499		
	/organism="uncultured bacterium"		
	/mol_type="genomic DNA"		

ORIGIN	Query Match	92.2%	Score 1343.2	DB 3	Length 1499
	Best Local Similarity	95.5%	Pred. No. 4.8e-14		
	Matches 1394	Conservative 0	Mismatches 63	Indels 3	Gaps 1
rRNA					
					/isolation_source="uncultured clone from polychlorinated dioxin dechlorinating microcosm" /db_xref="taxon:77133" /clone="TSAQ22" /environmental_sample <1..>1499 /product="16S ribosomal RNA"
QY	1	ATTGAACGCTGCGGCATGCTTTTACACATGCAAGTCGAACGGCGACGACGATGCTTGAT	60		
DB	21	ATTGAACGCTGCGGCATGCTTTTACACATGCAAGTCGAACGGCGACGACGATGCTTGAC	80		
QY	61	CTGTGTGGGAATGCGCGGAGCGGGTGAATATCATTCGAAACGTATCCAGAAAGGGGGGTA	120		
DB	81	TTGTGTGGGAATGCGCGGAGCGGGTGAATATCATTCGAAACGTATCCAGAAAGGGGGGTA	140		
QY	121	ACGCATCGAAAGATGTGCTAATATACCGCATATACTCTAAGAGAAAGCAGGGGATCGAA	180		
DB	141	ACGCATCGAAAGATGTGCTAATATACCGCATATACTCTAAGAGAAAGCAGGGGATCGAAA	200		
QY	181	GACCTTGGCGCTTTTGGAGCGCGCCGATGTCTGATTAGCTAAGTTGGTGGGGTTAAGCCCTAC	240		
DB	201	GACCTTATGCTTTTGGAGCGCGCCGATGTCTGATTAGCTAAGTTGGTGGGGTTAATTCCTAC	260		
QY	241	CAAGGCGACATCATGTAATTTGTCGAAAGGACGACGACCACTGGGGACTGAGACACG	300		
DB	261	CAAGGCGACATCATGTAATTTGTCGAAAGGACGACGACCACTGGGGACTGAGACACG	320		
QY	301	GCCGACATCTCTTACCGGAGGACGACGATGGGGAATTTTGGACAAATGGGCGCAAGCCTGATC	360		
DB	321	GCCGACATCTCTTACCGGAGGACGACGATGGGGAATTTTGGACAAATGGGCGCAAGCCTGATC	380		
QY	361	CAGCAATGCCCGGTGATGGAAGAGGCTTTGGGTTGTAAAGCTCTTTCAGTCGAGAAAG	420		
DB	381	CAGCAATGCCCGGTGATGGAAGAGGCTTTGGGTTGTAAAGCTCTTTCAGTCGAGAAAG	440		
QY	421	AAAGGTTACGGTAATTAATTCGATCATCGATGACGATGACAGAAAGAACCGGCTAAC	480		
DB	441	AAAGGTTACGGTAATTAATTCGATCATCGATGACGATGACAGAAAGAACCGGCTAAC	500		
QY	481	TACGTGCAGCAGCCGCGGTATATCGTAGGCTGCAAGCCTTAATCGAATTAATTCGAGCGT	540		
DB	501	TACGTGCAGCAGCCGCGGTATATCGTAGGCTGCAAGCCTTAATCGAATTAATTCGAGCGT	560		
QY	541	AAAGGTCGCGAGCGCGCTTTGTAAAGTCAGATGTGAATATCCCGGGCTTAACTCGGAT	600		
DB	561	AAAGGTCGCGAGCGCGCTTTGTAAAGTCAGATGTGAATATCCCGGGCTTAACTCGGAT	620		
QY	601	TGCGTTTGAATATCAAGGCTTGAAGTGGGAGGAGGAGGAGTGAATATTCATGTGTAGCAG	660		
DB	621	TGCGTTTGAATATCAAGGCTTGAAGTGGGAGGAGGAGGAGTGAATATTCATGTGTAGCAG	680		
QY	661	TGAATATGCTAAGATATGGAAGAAATCATGATGGCGAAAGCAGCCTCTGGGTTAACACT	720		
DB	681	TGAATATGCTAAGATATGGAAGAAATCATGATGGCGAAAGCAGCCTCTGGGTTAACACT	740		
QY	721	GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAAGATATCCTGGTATGTCCAGGCC	780		
DB	741	GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAAGATATCCTGGTATGTCCAGGCC	800		
QY	781	CTAAACGATGTAACATGATGTTGGGCTTTTAAAGCTGTGTACGAAGCTAACCGGTA	840		
DB	801	CTAAACGATGTAACATGATGTTGGGCTTTTAAAGCTGTGTACGAAGCTAACCGGTA	860		
QY	841	AGTTGACCGCTGGGAGTACGGTCCGACGAATTAATACTCAAGGAATTTGACGGGACCC	900		
DB	861	AGTTGACCGCTGGGAGTACGGTCCGACGAATTAATACTCAAGGAATTTGACGGGACCC	920		







Db	913	GCACACGCGGTGGATTATATGAGATTAAATTGATGCAACGGGAAAAACCTTA-CTACCCCTT	971
Qy	961	GACACTGTAGCGAAATTTTCTAGAGATAGATTAGTCTTCGGGAAACGTTAACACAGGTGCTG	1020
Db	972	GACATGTAGGAAATTTTCTAGAGATAGATTAGTCTTC-GGAACGCTTAACACAGGTGCTG	1030
Qy	1021	CATGGCTGTGTCAGCTTCGTGTGCTGAGATGTTGGGTTAAAGTCCCCGAACGAGCGCAAC	1080
Db	1031	CATGGCTGTGTCAGCTTCGTGTGCTGAGATGTTGGGTTAAAGTCCCCGAACGAGCGCAAC	1090
Qy	1081	CTTGTCAATTATTCGCATATTGGTTGGGCACTTTATGAGACTGCGCGGTGCAAAACG	1140
Db	1091	CTTGTCAATTATTCGCATATTAGTTGGGCACTTTATGAGACTGCGCGGTGCAAAACG	1150
Qy	1141	GAGGAAGGTGGGATGACGTCGACGTCTCATGACCCTTATGGGTAGGGCTTCAACGTAA	1200
Db	1151	GAGGAAGGTGGGATGACGTCGACGTCTCATGACCCTTATGGGTAGGGCTTCAACGTAA	1210
Qy	1201	TACAAATGGCGCGTACAGAGGGTTGCCAACCCCGCAAGGGGAGCTTAATCTCAGAAAGCGCG	1260
Db	1211	TACAAATGGCGCGTACAGAGGGTTGCCAACCCCGCAAGGGGAGCTTAATCTCAGAAAGCGCG	1270
Qy	1261	TCCGTAATGCCGATTCGGAGTCTGCAACTCGACTCCGTGAAAGTCGGAATCGTAATATCGC	1320
Db	1271	TCCGTAATGCCGATTCGGAGTCTGCAACTCGACTCCGTGAAAGTCGTAATATATCGC	1330
Qy	1321	GGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTATACACACCGCCCTGCACCAT	1380
Db	1331	GGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTATACACCGCCCTGCACCAT	1390
Qy	1381	GGGATGGGGTTTCAACAGAGCAGGTAGTCTTAACCGTAAAGAGGGCGCTTGCCACGTTGA	1440
Db	1391	GGGATGGGGTTTCAACAGAGCAGGTAGTCTTAACCGTAAAGAGGGCGCTTGCCACGTTGA	1450
Qy	1441	GATTCAATACCTGGGGTG 1457	
Db	1451	GATTCAATACCTGGGGTG 1467	
RESULT 24			
LOCUS	AY343318	1534 bp	DNA linear ENV 01-AUG-2005
DEFINITION	Uncultured bacterium clone OSI 16S ribosomal RNA gene, partial sequence.		
ACCESSION	AY343318		
VERSION	AY343318.1	GI:37813565	
KEYWORDS	ENV.		
SOURCE	uncultured bacterium		
ORGANISM	Bacteria; environmental samples.		
REFERENCE	1 (bases 1 to 1534)		
AUTHORS	Sliekers,A.O., Haeyer,S., Schmid,M., Harhangi,H., Verwegen,K., Kuennen,J.G. and Jetten,M.S.M.		
TITLE	Nitrification and Anammox with Urea as the Energy Source		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1534)		
AUTHORS	Sliekers,A.O., Haeyer,S., Schmid,M., Harhangi,H., Verwegen,K., Kuennen,J.G. and Jetten,M.S.M.		
TITLE	Direct Submision		
JOURNAL	Submitted (16-JUN-2003) Biotechnology, Delft University of Technology, Julianalaan 67, Delft 2628BC, The Netherlands		
FEATURES	Location/Qualifiers		
source	1..1534		
	/organism="uncultured bacterium"		
	/mol_type="genomic DNA"		
	/isolation_source="batch reactor"		
	/db_xref="taxon:77133"		
	/clone="OS1"		
	/environmental_sample		
	<1..>1534		
	/product="16S ribosomal RNA"		
ORIGIN			

Query Match		92.0%;	Score 1340;	DB 3;	Length 1534;
Best Local Similarity		95.3%;	Pred. No. 5.3e-14;		
Matches 1332;		Conservative	0;	Mismatches 65;	Indels 3;
				Gaps	1;
Qy	1	ATTGAACGCTGGCGCGCATCTTTTACATGCATGCAAGTGAACCGCAGACACGGATCTTCGCAT	60		
Db	21	ATTGAACGCTGGCGCGCATCTTTTACATGCATGCAAGTGAACCGCAGACACGGATCTTCGCAC	80		
Qy	61	CTGGTGGCGAATGGCGCGAAGGGGTAGTATGCAATCGTAAAGTATCCAGAAAGAGGGGGGTA	120		
Db	81	CTGGTGGCGAATGGCGCGAAGGGGTAGTATGCAATCGTAAAGTATCCAGAAAGAGGGGGGTA	140		
Qy	121	ACGCATCGAAAGATGTGCTAATATCCGCATATACTCTAAGAGGAGAAACAGGGGATCGAAA	180		
Db	141	ACGCATCGAAAGATGTGCTAATATCCGCATATACTCTAAGAGGAGAAAGTGGGATCGAAA	200		
Qy	181	GACCTTGCCTTTTGGAGCGCGCGATGTCTGATTACTAGTTAGTTGGTGGGTAAAGGCTTAC	240		
Db	201	GACCTTATGCTTTTGGAGCGCGCGATGTCTGATTACTAGTTAGTTAGTTAGTTAGTTAGTTAC	260		
Qy	241	CAAGCGGAGATATAGTATGTTGGTCTGAGAGAGACACAGCACACTGGGGGCTAGACAGC	300		
Db	261	CAAGCGGAGATATAGTATGTTGGTCTGAGAGAGACACAGCACACTGGGGGCTAGACAGC	320		
Qy	301	GCCCAACCTCTACGGGAGGACAGCATGGGGGAATTTTGGCAATGGGCGCAAGCCTGATC	360		
Db	321	GCCCAACCTCTACGGGAGGACAGCATGGGGGAATTTTGGCAATGGGCGCAAGCCTGATC	380		
Qy	361	CAGCAATGCGCGCTGATGATGAAGAGCGCTTCGGGTGTAAAGCTCTTTCAGTCGAGAGA	420		
Db	381	CAGCAATGCGCGCTGATGATGAAGAGCGCTTCGGGTGTAAAGCTCTTTCAGTTGAGAGA	440		
Qy	421	AAAGGTTACGGTAATATATCTGATCTCATGACGGTATCGACAGAAAGACACCGGCTAAC	480		
Db	441	AAAGCGGTGACTAATATATCATGTTCTATCACGGTATCGACAGAAAGACACCGGCTAAC	500		
Qy	481	TACGTCCAGACACCGCGGTAAATACGATGGGTCAAGCGTTAATCGAATTTACTGGGCGT	540		
Db	501	TACGTCCAGACACCGCGGTAAATACGATGGGTCAAGCGTTAATCGAATTTACTGGGCGT	560		
Qy	541	AAAGGGTGCAGAGCGGCTTTGTAAATCAGATGTGAATCCCGGGCTTAACTCGGGAAT	600		
Db	561	AAAGGGTGCAGAGCGGCTTTGTAAATCAGATGTGAATCCCGGGCTTAACTCGGGAAT	620		
Qy	601	TGCGTTGAAACTACAGGCTTAGAGTGTGCAAGGAGGATGGAATTCATGTGTACAG	660		
Db	621	TGCGTTGAAACTACAGGCTTAGAGTGTGCAAGGAGGATGGAATTCATGTGTACAG	680		
Qy	661	TGAATTCGTAAGATATGGAAGAACATCGATGGCCAAAGCAGCCTCTGGGTTAAACT	720		
Db	681	TGAATTCGTAAGATATGGAAGAACATCGATGGCCAAAGCAGCCTCTGGGTTAAACT	740		
Qy	721	GAGCTCATGACGAAAGCGTGGGGAGCAACAGGATTTAGATACCCTGTAGTCAAGCC	780		
Db	741	GAGCTCATGACGAAAGCGTGGGGAGCAACAGGATTTAGATACCCTGTAGTCAAGCC	800		
Qy	781	CTAAACGATGTCACTAGTTGTTGGGCTTATTAGGCTTGTTAAAGAACTTAACGCGTGA	840		
Db	801	CTAAACGATGTCACTAGTTGTTGGGCTTATTAGGCTTGTTAAAGAACTTAACGCGGTGA	860		
Qy	841	AGTTGACCGCTGGGGAGTACGGTGCAGAACTTAAACTCAAGAAATTGACGGGGAGCC	900		
Db	861	AGTTGACCGCTGGGGAGTACGGTGCAGAACTTAAACTCAAGAAATTGACGGGGAGCC	920		
Qy	901	GCAACAACGGGTGATATGAGATTTAATGATGCAACGCGAATACTTACCTTACCTT	960		
Db	921	GCAACAACGGGTGATATGAGATTTAATGATGCAACGCGAATACTTACCTTACCTT	980		
Qy	961	GACATGTAGCGAATTTTCTAGAGATATAGTCTT---CGGAGACGTTAAACAGGTG	1017		
Db	981	GACATGTAGCGAATTTTCTAGAGATATAGTCTT---CGGAGAGGAACTGTAAACAGGTG	1040		
Qy	1018	CTGATGAGCTGTCTGACCTGCTGTCTGTGAGATTTGGGTTAAGTCCCGCAACGAGCGCA	1077		

Db 1041 CTGCATGGCTGTCGTCAGCTCTGTCGTGAGATGTTGGTTAATCCCGCAACGAGCGCA 1100  
Qy 1078 ACCCTTGTCATTAATTCGCCATTTGTTGGGCACTTTAATGAGACTGCGGGTGAACA 1137  
Db 1101 ACCCTTGTCATTAATTCGCCATTTGTTGGGCACTTTAATGAGACTGCGGGTGAATAA 1160  
Qy 1138 CCGGAGGAAGGTGGGGATGACGTCAAGTCTCATGGCCCTTATGGGTAGGGCTTCAACG 1197  
Db 1161 CCGGAGGAAGGTGGGGATGACGTCAAGTCTCATGGCCCTTATGGGTAGGGCTTCAACG 1220  
Qy 1198 TAATCAATGGCGCGTACAGAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAAGAAAGC 1257  
Db 1221 TAATCAATGGCGCGTACAGAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAAGAAAGC 1280  
Qy 1258 GCGTGTAGTCCGATTCGAGAGTCTGCACTCGAATCCGTGAATCGGAATCGCTAGTAT 1317  
Db 1281 GCGTGTAGTCCGATTCGAGAGTCTGCACTCGAATCCGTGAATCGGAATCGCTAGTAT 1340  
Qy 1318 CCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCACAC 1377  
Db 1341 CCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCACAC 1400  
Qy 1378 CATGGAGTGGGTTTCCACGAGAGCATGATCTAACCGTAAAGAGGGCGCTTGCCACGG 1437  
Db 1401 CATGGAGTGGGTTTCCACGAGAGCATGATCTAACCGTAAAGAGGGCGCTTGCCACGG 1460  
Qy 1438 TGAGATTGATGACTGGGGTG 1457  
Db 1461 TGAGATTGATGACTGGGGTG 1480

RESULT 25  
LOCUS AJ621027 1454 bp DNA linear BCT 08-JUN-2004  
DEFINITION Nitrosomonas sp. Is32 16S rRNA gene, isolate Is32.  
ACCESSION AJ621027 GI:40994846  
VERSION AJ621027.1 GI:40994846  
KEYWORDS 16S ribosomal RNA; 16S rRNA gene.  
SOURCE Nitrosomonas sp. Is32  
ORGANISM Nitrosomonas sp. Is32  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.

REFERENCE  
AUTHORS 1. Nielsen, S., Baer Gilielsen, M.J., Revsbech, N., Laanbroek, H.J. and Bollmann, A.  
TITLE Influence of heterotrophic bacteria on the activity and growth behavior of ammonia-oxidizing bacteria  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1454)  
AUTHORS Bollmann, A.  
TITLE Direct Submision  
JOURNAL Submitted (15-JAN-2004) Bollmann A., Department for Microbial Ecology, NIOO-KNAW Centre for Limnology, Rijksstratweg 6, 3631 AC Nieuwersluis, NETHERLANDS  
FEATURES  
source location/Qualifiers  
1..1454  
/organism="Nitrosomonas sp. Is32"  
/mol\_type="genomic DNA"  
/isolate="Is32"  
/isolation\_source="fresh water sediment"  
/db\_xref="taxon:261293"  
/country="Netherlands:Lake Dronlermeer"  
1..1454  
/gene="16S rRNA"  
1..1454  
/gene="16S rRNA"  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 91.7%; Score 1336.2; DB 1; Length 1454;  
Beet Local Similarity 95.4%; Pred. No. 6.4e-14;  
Matches 1387; Conservative 0; Mismatches 64; Indels 3; Gaps 1;

Qy 1 ATGGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAAGCGGACGACGATGCTTGAT 60  
Db 1 ATGGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAAGCGGACGACGATGCTTGAT 60  
Qy 61 CTGTGGCGAGTGGCGGACGGGGTAGTATGTCATCGAACGTATTCAGAAAGGGGGGTA 120  
Db 61 CTGTGGCGAGTGGCGGACGGGGTAGTATGTCATCGAACGTATTCAGAAAGGGGGGTA 120  
Qy 121 ACGCATGGAAGATGTGCTAATACCGCATATATCTTAAGAGGAAGAAAGCGGATTCGAAA 180  
Db 121 ACGCATGGAAGATGTGCTAATACCGCATATATCTTAAGAGGAAGAAAGCGGATTCGAAA 180  
Qy 181 GACCTTGCGCTTTTGGAGACGGGCCGATGTCATATTAAGCTATGGGGGTTAAAGGCTTAC 240  
Db 181 GACCTTGCGCTTTTGGAGACGGGCCGATGTCATATTAAGCTATGGGGGTTAAAGGCTTAC 240  
Qy 241 CAAGCGCAGATCAAGTATGCTGAGAGGAGCAGCCAGCATCTGGGACTGAGACAG 300  
Db 241 CAAGCGCAGATCAAGTATGCTGAGAGGAGCAGCCAGCATCTGGGACTGAGACAG 300  
Qy 301 GCCGACTCTCTACGGGAGGAGCAGCATGGGGAAATTTGGACATGGCGCAAGCTGATC 360  
Db 301 GCCGACTCTCTACGGGAGGAGCAGCATGGGGAAATTTGGACATGGCGCAAGCTGATC 360  
Qy 361 CAGCAATCCCGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
Db 361 CAGCAATCCCGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
Qy 421 AAAAGTTACGGTAAATATCGTGAATCGTGAATCGTGAATCGTGAATCGTGAATCGTGAAT 480  
Db 421 AAAAGTTACGGTAAATATCGTGAATCGTGAATCGTGAATCGTGAATCGTGAATCGTGAAT 480  
Qy 481 TAAGTCCGATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAATCA 540  
Db 481 TAAGTCCGATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAATCA 540  
Qy 541 AAAAGTCCGATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAAT 600  
Db 541 AAAAGTCCGATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAAT 600  
Qy 601 TGCGTTGAAATCTAAGGATGAGTGAAGGATGAGTGAAGGATGAGTGAAGGATGAGTGAAGG 660  
Db 601 TGCGTTGAAATCTAAGGATGAGTGAAGGATGAGTGAAGGATGAGTGAAGGATGAGTGAAGG 660  
Qy 661 TGAATGCGTGAAGTATGAGAAACATCGATGGCGAAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
Db 661 TGAATGCGTGAAGTATGAGAAACATCGATGGCGAAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
Qy 721 GACGCTCATGACGAAACCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
Db 721 GACGCTCATGACGAAACCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
Qy 781 GACGCTCATGACGAAACCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
Db 781 GACGCTCATGACGAAACCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
Qy 841 AGTTGACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
Db 841 AGTTGACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
Qy 901 GCAAGAGGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTAT 960  
Db 901 GCAAGAGGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTAT 960  
Qy 961 GACATGTAGGAAATTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Db 961 GACATGTAGGAAATTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Qy 1018 CTGCATGGCTGTCGTCAGCTCTGTCGTGAGATGTTGGGTTAAGTCCGCAACGAGCGCA 1077  
Db 1021 CTGCATGGCTGTCGTCAGCTCTGTCGTGAGATGTTGGGTTAAGTCCGCAACGAGCGCA 1080

QY	1078	ACCCGTGTCATTTAATGCCATCATTTGGTTGGGCACTTTAATGAGACTGCCGGTGACAA	1137
Db	1081	ACCCGTGTCATTTAATGCCATCATTTAGTTGGGCACTTTAATGAGACTGCCGGTGATAA	1140
QY	1138	CCGAGAGAAAGGTGGGGAGAGCGTCAAGCTCATGGCCCTTAATAGGTAGGGCTTCACACG	1197
Db	1141	CCGAGAGAAAGGTGGGGAGAGCGTCAAGCTCATGGCCCTTAATAGGTAGGGCTTCACACG	1200
QY	1198	TAATACATATGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGGAGCTAATCTCAAGAAAGC	1257
Db	1201	TAATACATATGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGGAGCTAATCTCAAGAAAGC	1260
QY	1258	GCGTGTAGTCCGAGTCCGAGTCTTGCAACTGCACTCCGTGAATGCGGAATGCTAAT	1317
Db	1261	GCGTGTAGTCCGAGTCCGAGTCTTGCAACTGCACTCCGTGAATGCGGAATGCTAAT	1320
QY	1318	CGCGAATACAGATGTCGCGGTGAATATGCTCCGGGTCTTGTATACACACCGCCCGTACAC	1377
Db	1321	CGCGAATACAGATGTCGCGGTGAATATGCTCCGGGTCTTGTATACACACCGCCCGTACAC	1380
QY	1378	CATGGAGATGGGTTTACACAGAGCAGGTAGTCTAACCGTAAGAGGCGGTTGCCACGG	1437
Db	1381	CATGGAGATGGGTTTACACAGAGCAGGTAGTCTAACCGTAAGAGGCGGTTGCCACGG	1440
QY	1438	TGAGATTCATGACT 1451	
Db	1441	TGAGATTCATGACT 1454	
RESULT 26	AY123811	1499 bp DNA linear BCT 23-SEP-2003	
LOCUS	AY123811	Nitrosomonas sp. Nm59 16S ribosomal RNA gene, partial sequence.	
DEFINITION	AY123811	GI:24474440	
ACCESSION	AY123811.1	GI:24474440	
VERSION	AY123811.1	GI:24474440	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE		</	

Db	63	TNCGTGCAGATGGCGAA	CGGGTGAGTAATGCAATCGGAACGTGTCTTAAAGTGGGAGATA	122
Oy	121	ACCGATGGAAAGATNGCTAA	TACCCGATATCTTAAGGAGAAAGACGGGGATCGMAA	180
Db	123	ACGGATGAAAGATGCTAATAC	CGCATATCTCTCAGAGGAAAGTAGGGGATCGMAA	182
Oy	181	GACCTTCGCGTTTGGAGCG	CGCGATGCTGATTAAGTAGTGTGGGTTAAAGCCCTAC	240
Db	183	GACCTTATGCTTTAAGAGCG	CGGCGATGTCTGATTAAGTAGTGTGGGTTAAAGCCCTAC	242
Oy	241	CAAGGCGACATCATGTA	AGTTGGTCTGAGAGAGACCAAGCCACATCTGGGACTGAGACAG	300
Db	243	CAAGGCGACATCATGTA	AGTTGGTCTGAGAGAGACCAAGCCACATCTGGGACTGAGACAG	302
Oy	301	GCCCAAGCTCCATAGGGAG	CGACAGTGGGGAAATTTTGGACAATGGGCGCAAGCCTGATC	360
Db	303	GCCCAAGCTCCATAGGGAG	CGACAGTGGGGAAATTTTGGACAATGGGCGCAAGCCTGATC	362
Oy	361	CAGCAATGCCCGCTGATG	AAGAAGCCCTTCGGGTTGTAAGCTCTTTCAGTCGAGAGA	420
Db	363	CAGCAATGCCCGCTGATG	AAGAAGCCCTTCGGGTTGTAAGCTCTTTCAGTTGAGAGA	422
Oy	421	AAAGGTTACGGTAAATTA	TGCTGACTCATGACGCTATCGACAGAAAGACACGGGCTTAAC	480
Db	423	AAAAATGTGACTAATATCA	CAATTGATGACGGTATCAACAGAGAAAGACACGGGCTTAAC	482
Oy	481	TACGTGCCAGAGCGCGGT	TAATCGTAAGGTGCAACGGTAAATCGGAATTAACCTGAGGCGT	540
Db	483	TACGTGCCAGAGCGCGGT	TAATCGTAAGGTGCAACGGTAAATCGGAATTAACCTGAGGCGT	542
Oy	541	AAAGGGTSCGACAGCGG	CTTGTGTAAGTCAGATGTGAATCCCGGGGCTTAACCTGGGAAT	600
Db	543	AAAGGGTSCGACAGCGG	CTTGTGTAAGTCAGATGTGAATCCCGGGGCTTAACCTGGGAAT	602
Oy	601	TGCGTTTGAATCTA	CAGAGCTTAAGTGTGGCAGAGGGAGGTGGAAATTCATGTGCTAACAG	660
Db	603	TGCGTTTGAATCTA	CAGAGCTTAAGTGTGGCAGAGGGAGGTGGAAATTCATGTGCTAACAG	662
Oy	661	TGAAATGGGTGAGATAT	AGAAAGCAATCGATGSCAGAAAGGACGCTCCTGGGTTAAACACT	720
Db	663	TGAAATGGGTGAGATAT	AGAAAGCAATCGATGSCAGAAAGGACGCTCCTGGGTTAAACACT	722
Oy	721	GACGCTCATGACAGAA	AGCCTGGGAGCAAAACAGAAATTAAGTACCTTGGTATGTCACAGCC	780
Db	723	GACGCTCATGACAGAA	AGCCTGGGAGCAAAACAGAAATTAAGTATCCCTGGTATGTCACAGCC	782
Oy	781	CTAAACGATGTC	AACTAAGTGTGGGCTTATTAAGGCTTGGTAAACGAAGCTTAACGCGTGA	840
Db	783	CTAAACGATGTC	AACTAAGTGTGGGCTTATTAAGGCTTGGTAAACGAAGCTTAACGCGTGA	842
Oy	841	AGTTGACCGGCTGGGAG	ATAGCGTCCGAAGTTAAACTCAAAAGAAATGACGCGGAGACC	900
Db	843	AGTTGACCGGCTGGGAG	ATAGCGTCCGAAGTTAAACTCAAAAGAAATGACGCGGAGACC	902
Oy	901	GCACAAGCGGTGATTA	TATGTGGAATTAATTCGATGCAACGCGAAAAACCTTAACCTTACCCTT	960
Db	903	GCACAAGCGGTGATTA	TATGTGGAATTAATTCGATGCAACGCGAAAAACCTTAACCTTACCCTT	962
Oy	961	GACATGTAGCGAA	TTTCTAAGATAGATTAGTGC---CTTCGGGAGCGCTTAACACAGGTG	1017
Db	963	GACATGTAGCGAA	TTTCTAAGATAGATTAGTGC---CTTCGGGAGCGCTTAACACAGGTG	1022
Oy	1018	CTGCATAGCGTGTG	CGACGTGTGCGAGATGTGGGTTAAGTCCCGCAACGAGCGCA	1077
Db	1023	CTGCATAGCGTGTG	CGACGTGTGCGAGATGTGGGTTAAGTCCCGCAACGAGCGCA	1082
Oy	1078	ACCCCTTGATTAAT	TGCGCATATTTGGTGGGCACTTTAATGACACTGCGGTGACAAA	1137
Db	1083	ACCCCTTGATTAAT	TGCGCATATTTGGTGGGCACTTTAATGACACTGCGGTGACAAA	1142
Oy	1138	CCGAGAGAA	GTGGGATGACGTCAAGTCTCTATGGCCCTTAATGGGTAGGGCTTACACG	1197

Db 1143 CCGAGGAAGTGGGATGACGTCAAGTCTCATGSCCTTATGGGTAGGGCTGCACG 1202  
Qy 1198 TAAATCAATGGCGCTGACAGAGGTTTCCCAACCCGAGGGGAGCTAATCTCAGAAAGC 1257  
Db 1203 TAAATCAATGGCGCTGACAGAGGTTTCCCAACCCGAGGGGAGCTAATCTCAGAAAGC 1262  
Qy 1258 GGGTGTAGTCCGGATCCGAGTCTGCACTCCGTAAGTGGGAATGGCTAGTAT 1317  
Db 1263 GGGTGTAGTCCGGATCCGAGTCTGCACTCCGTAAGTGGGAATGGCTAGTAT 1322  
Qy 1318 CGCGATCAGCATGTCCGCGTAATACGTCCCGGCTCTTGTACACCCGCGCTCACAC 1377  
Db 1323 CGCGATCAGCATGTCCGCGTAATACGTCCCGGCTCTTGTACACCCGCGCTCACAC 1382  
Qy 1378 CATGGAGTGGGTTTCCAGGAAGCATGTCTTAAACCTTAAGAGAGGCGCTTGCACGG 1437  
Db 1383 CATGGAGTGGGTTTCCAGGAAGCATGTCTTAAACCTTAAGAGAGGCGCTTGCACGG 1442  
Qy 1438 TGAGATTCATGACTGGGGTGTG 1457  
Db 1443 TGAGATTCATGACTGGGGTGTG 1462

RESULT 27  
AY123810 1499 bp DNA linear BCT 23-SRP-2003  
LOCUS Nitrosomonas sp. Nm47 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY123810  
ACCESSION AY123810.1 GI:24474439  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Nitrosomonas sp. Nm47  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1 (bases 1 to 1499)  
Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roseer,A. and  
Koops,H.P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
JOURNAL PUBMED 13130037  
REFERENCE 2 (bases 1 to 1499)  
Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Roseer,A. and  
Wagner,M.  
TITLE Direct Submission  
Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
JOURNAL  
AUTHORS  
FEATURES  
source location/Qualifiers  
1..1499  
/organism="Nitrosomonas sp. Nm47"  
/mol\_type="genomic DNA"  
/strain="Nm47"  
/db\_xref="taxon:200128"  
<1..>1499  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 91.4%; Score 1311; DB 1; Length 1499;  
Best Local Similarity 94.9%; Pred. No. 7, 4e-14;  
Matches 186; Conservative 0; Mismatches 71; Indels 3; Gaps 1;

Qy 1 ATTGAACGCTGGCGGATGCTTTACATGCAAGTCGAAACGCGACGAGATGCTTGGCT 60  
Db 3 ATTGAACGCTGGCGGATGCTTTACATGCAAGTCGAAACGCGACGAGATGCTTGGCT 62  
Qy 61 CTGGTGGGAGTGGGAGACGGGTGATATGATCGGAACGATCAGAAAGAGGGGGGAT 120  
Db 63 CNGGTGGGAGTGGGAGACGGGTGATATGATCGGAACGATGCTTGAAGTGGGGGAT 122  
Qy 121 ACGCATCGAAAGATGTGCTAATACGATATATCTTAAGAGAGAAAGCAGGGGATCGAA 180  
Db 123 ACGCATCGAAAGATGTGCTAATACGATATATCTTAAGAGAGAAAGTGGGGATCGTA 182

Qy 181 GACCTTGGCTTTTGGAGGCGCCGATGCTGATTAAGTATGGTGGGTTAAAGCCTTAC 240  
Db 183 GACCTTATGCTTTTGGAGGCGCCGATGCTGATTAAGTATGGTGGGTTAAAGCCTTAC 242  
Qy 241 CAAGCGCATGATGATGATGCTGAGAGAGACACACGACACTGGGACTGAGACAG 300  
Db 243 CAAGCGCATGATGATGATGCTGAGAGAGACACACGACACTGGGACTGAGACAG 302  
Qy 301 GCCCAACTCTTACGGGAGGACAGATGGGGAATTTTGGACAAATGGGCGCAAGCTTATC 360  
Db 303 GCCCAACTCTTACGGGAGGACAGATGGGGAATTTTGGACAAATGGGCGCAAGCTTATC 362  
Qy 361 CAGCAATCCCGCTGAGTGAAGAAAGCCCTTGGGGTGAAGCTTTCACTGAGGAAGA 420  
Db 363 CAGCAATCCCGCTGAGTGAAGAAAGCCCTTGGGGTGAAGCTTTCACTGAGGAAGA 422  
Qy 421 AAAGGTTACGGTAAATATCGTGACTCATGACGGTATCGACAGAAAGACACCGCTAAC 480  
Db 423 AAAGGTTACGGTAAATATCGGATTTATGACGGTATCGACAGAAAGACACCGCTAAC 482  
Qy 481 TACGTCCAGACGCCCGGCTAATACGTATGGGTGCAAGCCTTATCGGAATTCAGGGCT 540  
Db 483 TACGTCCAGACGCCCGGCTAATACGTATGGGTGCAAGCCTTATCGGAATTCAGGGCT 542  
Qy 541 AAAGGTTACGGAGGCGCTTGTAAATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 543 AAAGGTTACGGAGGCGCTTGTAAATGATGATGATGATGATGATGATGATGATGATGAT 602  
Qy 601 TCGCTTTGAAACTACAGAGCTAGAGTGGGAGAGAGAGTGGAAATTCATGTGTAGAG 660  
Db 603 TCGCTTTGAAACTAGAGTGGGAGAGAGAGTGGAAATTCATGTGTAGAGAG 662  
Qy 661 TGAATGTGTAGATATGGAAGAACTCATGATGCGAAGGACGCTCCTGGGTTAACACT 720  
Db 663 TGAATGTGTAGATATGGAAGAACTCATGATGCGAAGGACGCTCCTGGGTTAGCACT 722  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGGAGCAAGGATTAATGATGATGATGATGATGATGAT 780  
Db 723 GACGCTCATGACGAAAGCGTGGGAGGAGCAAGGATTAATGATGATGATGATGATGATGAT 782  
Qy 781 CTAACGATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
Db 783 CTAACGATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842  
Qy 841 AGTTGACCGCTGGGAGTACGCTGCGCAAGTTAAATCTCAAAGAAATTTGACGGGAGACC 900  
Db 843 AGTTGACCGCTGGGAGTACGCTGCGCAAGTTAAATCTCAAAGAAATTTGACGGGAGACC 902  
Qy 901 GCAAGAGGAGTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Db 903 GCAAGAGGAGTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962  
Qy 961 GACATGTAGCAATTTTCTAGAGATAGATAGTATGCT--TCGGAAACGCTAACACAGGTG 1017  
Db 963 GACATGTAGCAATTTTCTAGAGATAGATAGTATGCT--TCGGAAACGCTAACACAGGTG 1022  
Qy 1018 CTGCATGCTGCTGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
Db 1023 CTGCATGCTGCTGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082  
Qy 1078 ACCCTGTGCTAATTTGCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
Db 1083 ACCCTGTGCTAATTTGCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142  
Qy 1138 CCGAGGAAGTGGGAGTGAAGTCTCATAGCCCTTATGGGTAGGGCTTCAACG 1197  
Db 1143 CCGAGGAAGTGGGAGTGAAGTCTCATAGCCCTTATGGGTAGGGCTTCAACG 1202  
Qy 1198 TAAATCAATGGCGCTGACAGAGGTTGCGAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1257  
Db 1203 TAAATCAATGGCGCTGACAGAGGTTGCGAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1262

QY 1258 GCGTGTAGTCGGATTCGGAGTCTGCACTGCTCCGTGAAGTGGAACTCCCTAGTAT 1317  
| | | | |  
Db 1263 GCGTGTAGTCGGATTCGGAGTCTGCACTGCTCCGTGAAGTGGAACTCCCTAGTAT 1322  
| | | | |  
QY 1318 CCGCGATCAGCATGTCTCGCGGTAAATACCTTCCCGGCTCTTGACACACCGCCGACAC 1377  
| | | | |  
Db 1323 CCGCGATCAGCATGTCTCGCGGTAAATACCTTCCCGGCTCTTGACACACCGCCGACAC 1382  
| | | | |  
QY 1378 CATGGAGTGGGTTTCCACGAGAGCAGTAGTCTAAACCGTAAGAGAGCGCTTGGCCACG 1437  
| | | | |  
Db 1383 CATGGAGTGGGTTTCCACGAGAGCAGTAGTCTAAACCGTAAGAGAGCGCTTGGCCACG 1442  
| | | | |  
QY 1438 TGAGATTCACTGACTGGGCTG 1457  
| | | | |  
Db 1443 TGAGATTCACTGACTGGGCTG 1462  
| | | | |  
RESULT 28  
AF272422 1470 bp DNA linear BCT 05-DEC-2000  
LOCUS AF272422 Nitrosomonas oligotropha 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF272422  
ACCESSION AF272422 GI:11545282  
VERSION AF272422.1  
KEYWORDS  
SOURCE Nitrosomonas oligotropha  
ORGANISM Nitrosomonas oligotropha  
Bacteria: Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1 (bases 1 to 1470)  
AUTHORS Purkhold,U., Pommerehning-Rosser,A., Juretschko,S., Schmid,M.C.,  
Koops,H.P. and Wagner,M.  
TITLE Phylogeny of all recognized species of ammonia oxidizers based on  
molecular diversity surveys  
JOURNAL Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)  
PUBMED 11097916  
REFERENCE 2 (bases 1 to 1470)  
AUTHORS Purkhold,U., Pommerehning-Rosser,A., Juretschko,S., Schmid,M.C.,  
Koops,H.P. and Wagner,M.  
TITLE Direct Submission  
JOURNAL Submitted (26-MAY-2000) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source 1..1470  
/organism="Nitrosomonas oligotropha"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:42354"  
RNA <1..>1470  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 90.8%; Score 1323; DB 1; Length 1470;  
Best Local Similarity 95.0%; Pred. No. 1e-13;  
Matches 1375; Conservative 0; Mismatches 69; Indels 3; Gaps 1;  
QY 5 AACGCTGGCGGCAATGCTTTACACATGCAATGCAACGCGACGCGATGCTTGATCTGG 64  
| | | | |  
Db 24 AACGCTGGCGGCAATGCTTTACACATGCAATGCAACGCGACGCGATGCTTGATCTGG 83  
| | | | |  
QY 65 TGGCGAGTGGCGGAGCGGGTGAATGATCATCGGAACGTATCCAGAAAGGGGGTAAACGC 124  
| | | | |  
Db 84 TGGCGAGTGGCGGAGCGGGTGAATGATCATCGGAACGTATCCAGAAAGGGGGTAAACGC 143  
| | | | |  
QY 125 ATCGAAAGATGTGCTAATACCGCATATCTCTTAAGAGGAAAGAGGGGATCGAAAGCC 184  
| | | | |  
Db 144 ATCGAAAGATGTGCTAATACCGCATATCTCTTAAGAGGAAAGAGGGGATCGAAAGCC 203  
| | | | |  
QY 185 TTGGGCTTTTGGAGCGGCGGATGCTGATTAGTAGTTGGTGGGTTAAAGGCTTACCAAG 244  
| | | | |  
Db 204 TTATGCTTTTGGAGCGGCGGATGCTGATTAGTAGTTGGTGGGTTAAAGGCTTACCAAG 263  
| | | | |  
QY 245 GCGACGATCACTAGTGTGCTGAGAGGACGACGCACTGGAGTGAAGACGCGCC 304  
| | | | |

Db 264 GCAACGATCAGTAGTGTGCTGAGAGCGGCCACGACCACTGGGACTGAGACAGCGCC 323  
| | | | |  
QY 305 AGACTCTTACGGAGAGCGAGCTGGGAAATTTTGGACAAATGGGGCGAACCGTATCCAGC 364  
| | | | |  
Db 324 AGACTCTTACGGAGAGCGAGCTGGGAAATTTTGGACAAATGGGGCGAACCGTATCCAGC 383  
| | | | |  
QY 365 AATGCGGCGTGAAGTAAGAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGTAAGAAAAG 424  
| | | | |  
Db 384 AATGCGGCGTGAAGTAAGAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGTAAGAAAAG 443  
| | | | |  
QY 425 GTTACGTAATTAATCGTACTCATGACGTATGACAGAGAACACCGGCTTAATCTAGC 484  
| | | | |  
Db 444 ATTNNAGCTAATTAATCATTAATTCATGACGTATCAACAGAGAACACCGGCTTAATCTAGC 503  
| | | | |  
QY 485 TGCCAGACGCGCGGTAATACGTAGGCTGCAAGCTTAAATCGGAATTAATCGGCGGTAAG 544  
| | | | |  
Db 504 TGCCAGACGCGCGGTAATACGTAGGCTGCAAGCTTAAATCGGAATTAATCGGCGGTAAG 563  
| | | | |  
QY 545 GGTGGCGAGGCGGCTTTGTAAGTCAATGATGTAATCCCGGCTTAACTGGGAATTCGC 604  
| | | | |  
Db 564 GGTGGCGAGGCGGCTTTGTAAGTCAATGATGTAATCCCGGCTTAACTGGGAATTCGC 623  
| | | | |  
QY 605 TTTGAAACTACAAAGCTAGAGTGGCGAGAGAGGTGAATTCATGTTGTAGCACTGAA 664  
| | | | |  
Db 624 TTTGAAACTACAAAGCTAGAGTGGCGAGAGAGGTGAATTCATGTTGTAGCACTGAA 683  
| | | | |  
QY 665 ATGCGTAGATATGGAAGAAATATGATGGCGAAGGCAAGCTTCCGGGTTAACTAGTACG 724  
| | | | |  
Db 684 ATGCGTAGATATGGAAGAAATATGATGGCGAAGGCAAGCTTCCGGGTTAGAGCTAGAG 743  
| | | | |  
QY 725 CTCAGGACAGAAAGGTGGGGAGCAAAACAGATTAGATACCTGTAGTCCAGCCCTTA 784  
| | | | |  
Db 744 CTCAGGACAGAAAGGTGGGGAGCAAAACAGATTAGATACCTGTAGTCCAGCCCTTA 803  
| | | | |  
QY 785 ACGATGTCAACTAGTGTGGCTTATTAAGCTTGTGTAACGAAGCTAACGCGTGAATT 844  
| | | | |  
Db 804 ACGATGTCAACTAGTGTGGCTTATTAAGCTTGTGTAACGAAGCTAACGCGTGAATT 863  
| | | | |  
QY 845 GACCGCTGGAGATACGCTCGAGATTTAAACTAAAGAAATTCAGCGGGACCGGCAC 904  
| | | | |  
Db 864 GACCGCTGGAGATACGCTCGAGATTTAAACTAAAGAAATTCAGCGGGACCGGCAC 923  
| | | | |  
QY 905 AACCGGTGATTAATGAGATTAATTCAGTGAACGCGAAACCTTAACCTACCTTGACA 964  
| | | | |  
Db 924 AACCGGTGATTAATGAGATTAATTCAGTGAACGCGAAACCTTAACCTACCTTGACA 983  
| | | | |  
QY 965 TGTAGCGAATTTTCTAGATAGATTAGTGC--TTGGGAAACGCTAACACAGTGTCTGC 1021  
| | | | |  
Db 984 TGTAGCGAATTTTCTAGATAGATTAGTGC--TTGGGAAACGCTAACACAGTGTCTGC 1043  
| | | | |  
QY 1022 ATGGCTGTGTCAGCTGCTGTCGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACC 1081  
| | | | |  
Db 1044 ATGGCTGTGTCAGCTGCTGTCGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACC 1103  
| | | | |  
QY 1082 TTGTCAATTAATTCGCAATTTGGTGGGCACTTAATGAGATGCGCGGTGAACAACCG 1141  
| | | | |  
Db 1104 TTGTCAATTAATTCGCAATTTGGTGGGCACTTAATGAGATGCGCGGTGAACAACCG 1163  
| | | | |  
QY 1142 AGGAAGTGGAGATGACGTCAATGCTCAATGCGCTTATGGTATGGGCTTCAACGTAAT 1201  
| | | | |  
Db 1164 AGGAAGTGGAGATGACGTCAATGCTCAATGCGCTTATGGTATGGGCTTCAACGTAAT 1223  
| | | | |  
QY 1202 ACAATGGCGGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGCGT 1261  
| | | | |  
Db 1224 ACAATGGCGGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGCGT 1283  
| | | | |  
QY 1262 CGTAGTCCGAGATGAGATCTGCAACTCGACTCGGTAAGCGGAATCGCTAGTAATCGCG 1321  
| | | | |  
Db 1284 CGTAGTCCGAGATGAGATCTGCAACTCGACTCGGTAAGCGGAATCGCTAGTAATCGCG 1343  
| | | | |  
QY 1322 GATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGCTCACACCATG 1381  
| | | | |  
Db 1344 GATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGCTCACACCATG 1403  
| | | | |

OY		1382	GGAGTGGGTTTCACCAAGAAGCATGTACTAACCGTAAGAGAGGGGGCTTGCCAGCGTAG	1441
Db		1404	TGAGTGGGTTTACCAGAAAGCAGTAGTCTAACCATTAGAGAGGGCGCTTGANNACGGTAG	1463
OY		1442	ATTCATG	1448
Db		1464	GTTCAATG	1470
RESULT_29				
LOCUS	AB000699		1501 bp	DNA linear BCT 06-MAR-2001
DEFINITION	Nitrosomonas sp. AL212 gene for 16S rRNA, partial sequence.			
ACCESSION	AB000699			
VERSION	AB000699.1		GI:3107908	
KEYWORDS				
SOURCE				
ORGANISM	Nitrosomonas sp. AL212 Nitrosomonas sp. AL212 Bacteria; Proteobacteria; Betaproteobacteria; Nitrospirae; Nitrospirobacteraceae; Nitrospira			
REFERENCE				
AUTHORS	Suwa, Y., Sumino, T. and Noto, K.			
TITLE	Phylogenetic relationships of activated sludge isolates of ammonia oxidizers with different sensitivities to ammonium sulfate			
JOURNAL	J. Gen. Appl. Microbiol. 43, 373-379 (1997)			
REFERENCE	2. (bases 1 to 1501) Suwa, Y. Direct Submission Submitted (24-JAN-1997) Yuichi Suwa, National Institute for Resources & Environment, Ecological Chemistry and Microbiology Division; 16-3 Onogawa, Tsukuba, Ibaraki 305, Japan (E-mail:suwa@nre.go.jp Tel:+81-298-58-8310, Fax:+81-298-58-8309)			
FEATURES				
source	1..1501 /organism="Nitrosomonas sp. AL212" /mol_type="genomic DNA" /isolate="AL212" /db_xref="taxon:153948" "1..>1501 /product="16S ribosomal RNA"			
ORIGIN				
Query Match	90.8%; Score 1322.4; DB 1; Length 1501;			
Best Local Similarity	94.6%; Pred. No. 1e-13; 76; Indels 3; Gaps 1			
Matches 1381; Conservative 0; Mismatches 1381				
OY		1	ATTGAACCTGGCGGCATGCTTTACACATGCAGATGGAACGGCAGACGGATGCTTGCA	60
Db		21	ATTGAACCTGGCGGCATGCTTTACACATGCAGATGGAACGGCAGACGGATGCTTGCA	80
OY		61	CTGTGTGGCGAGTGGCGGACGGGTGAGTAATGCATCGGAACCTATCCAGAAAGGGGGGTA	120
Db		81	CCGGTGGCGAGTGGCGGACGGGTGAGTAATGCATCGGAACCTATCCAGAAAGGGGGGTA	140
OY		121	ACGCATCGAAGATGTCATAATACCGCATATCTCTAAGAGAGAAACGAGGATCTGA	180
Db		141	ACGCATCGAAGATGTCATAATACCGCATATCTCTAAGAGAGAAACGAGGATCTGA	200
OY		181	GACCTTGGCGCTTTGGAGCGGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGCCTTAC	240
Db		201	GACCTTGGCGCTTTGGAGCGGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGCCTTAC	260
OY		241	CAGGCGAGCATCACTAGTTGTCTGAGAGACGACCAACCACTGGGATCTGAGACACG	300
Db		261	CAGGCGAGCATCACTAGTTGTCTGAGAGACGACCAACCACTGGGATCTGAGACACG	320
OY		301	GCCCAGATCTCCATACGGGAGGACGACGTGGGAAATTTGGACATGGGCGCAAGCCTGATC	360
Db		321	GCCCAGATCTCCATACGGGAGGACGACGTGGGAAATTTGGACATGGGCGCAAGCCTGATC	380
OY		361	CAGCATGCGCGGTGATGAGAAAGCCCTTGGGGTTGTAAAGCTCTTTCATGTCAGAAACA	420

Db	381	CAGCAATGCCCGTGAATGGAAGAACGCCCTTCGGGTTGTAAGCTTTCAAGTTAGAAAGA	440
Oy	421	AAAGGTTACGGTAAATATATCGTGACTCATGAACGGTATTCGACAGAAAGACACCGGCTAAC	480
Db	441	AAAAATTCCTGGCTAATATCTGGAATTCATGACCGGTATCAACAGAAAGAACCCGGCTAAC	500
Oy	481	TACGTGCCAGAGCGCGGGTAAATACGTAAGGGTGTGAACGGTTAATCCGAATTACTGGGGCGT	540
Db	501	TACGTGCCAGAGCGCGGGTAAATACGTAAGGGTGTGACGGTTAATCCGAATTACTGGGGCGT	560
Oy	541	AAAGGGTACCGAGCGCGGCTTTGTGAAGTCAGATGTGAATCCCGGGGCTTAACTGGGAAT	600
Db	561	AAAGGGTACCGAGCGGCTTTGTGAAGTCAGATGTGAATCCCGGGGCTTAACTGGGAAT	620
Oy	601	TGCGTTTGAAACTCAAGGCTTAGAGTGATGCGCAGAGGGAGGTGGAATTCATGTGTACAG	660
Db	621	TGCGTTTGAAACTCAAGGCTTAGAGTGATGCGCAGAGGGGGGTGGAATTCATGTGTACAG	680
Oy	661	TGAAATGCGTGAAGATATGAAAGAACATCCATGGCGAAGGCAAGCCTCTGGGTTAACACT	720
Db	681	TGAAATGCGTGAAGATATGAAAGAACATCCATGGCGAAGGCAAGCCTCTGGGTTAACACT	740
Oy	721	GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGAAATTAGTACCTCGTGTATGCCAGCC	780
Db	741	GACACTACAGGACAGAAAGCGTGGGAGCAAAACAGAAATTAGTATACCTCGTGTATGCCAGCC	800
Oy	781	CTAAACGATGTCAACTAGTGTGTTGGGCTTATTAAGGCTTGGTAAAGAACTTAACGCGTGA	840
Db	801	CTAAACGATGTCAACTAGTGTGTTGGGCTTATTAAGGCTTGGTAAAGAACTTAACGCGTGA	860
Oy	841	AGTTGACCGGCTTGGGAGTACGATCGTGCAGATTTAAACTCAAAGAAATTGACGGGGACCC	900
Db	861	AGTTGACCGGCTTGGGAGTACGATCGTGCAGATTTAAACTCAAAGAAATTGACGGGGACCC	920
Oy	901	GCACAAGGGGTGGAATTAATGGAATTAATTGCATGCAACGCGAAACCTTACTACACCTT	960
Db	921	GCACAAGGGGTGGAATTAATGGAATTAATTGCATGCAACGCGAAACCTTACTACACCTT	980
Oy	961	GACATGTAGCCAAATTTTCTAGAGATAGATTAGTCT--TCGGGAACGCTTAACACAGGTG	1017
Db	981	GACATGTAGCAAAAATTTCAAGATAGATTAGTCTCGAAGAAAGAAATCTGAACACAGGTG	1040
Oy	1018	CTGCATGGCTGTCTGCACTGTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA	1077
Db	1041	CTGCATGGCTGTCTGCACTGTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA	1100
Oy	1078	AACCTTGTCATTAATTCGCATCATTTGGGTTGGGCACTTTAATGACATGCGCGGTGACAA	1137
Db	1101	AACCTTGTCATTAATTCGCACATTAATGATTGGCACTTTAATGACATGCGCGGTGATMA	1160
Oy	1138	CCGAGGAAAGGTGGGGAATGACGTCAAGTCTCATAGGCCCTTATAGGGGTTACAGCG	1197
Db	1161	CCGAGGAAAGGTGGGGAATGACGTCAAGTCTCATAGGCCCTTATAGGGGTTACAGCG	1220
Oy	1198	TAAATCAATGGCGCGTACAGAGGTTGCCAACCCGCGAAGGGGAGCTAATTCAGAAAGC	1257
Db	1221	TAAATCAATGGCGCGTACAGAGGTTGCCAACCTGTAGAGGGGAGCTAATTCAGAAAGC	1280
Oy	1258	GCGTCGTAGTCCGATCGGAATCTGCAACTCGATCCGTGAAAGTCGGAATGCTAGTAAAT	1317
Db	1281	GCGTCGTAGTCCGATCGGAATCTGCAACTCGATCCGTGAAAGTCGGAATGCTAGTAAAT	1340
Oy	1318	CGCGGATCAGCATGTCCGCGTGAATACGTTCCCGGTCCTTGTATACACACCGCCGCTCAC	1377
Db	1341	CGCGGATCAGCATGTCCGCGTGAATACGTTCCCGGTCCTTGTATACACACCGCCGCTCAC	1400
Oy	1378	CATGGAGTGGGTTTACACAGAACAGGTAGTCTTAACCGTAAGAGAGGCGCTTGCACCG	1437
Db	1401	CATGGAGTGGGTTTACACAGAACAGGTAGTCTTAACCGTAAGAGAGGCGCTTGCACCG	1460
Oy	1438	TGAATTTATATGACTGGGGTG 1457	
Db	1461	TGAAGTTATATGACTGGGGTG 1480	

RESULT 30  
LOCUS AF272423 1526 bp DNA linear BCT 05-DEC-2000  
DEFINITION Nitrosomonas cryotolerans 16S ribosomal RNA gene, partial sequence.  
ACCESSION AF272423  
VERSION AF272423.1 GI:11545283  
KEYWORDS  
SOURCE Nitrosomonas cryotolerans  
ORGANISM Nitrosomonas cryotolerans  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
1 (bases 1 to 1526)  
Purthold,U., Pommerening-Roseer,A., Juretschko,S., Schmid,M.C.,  
Koops,H.P. and Wagner,M.  
Phylogeny of all recognized species of ammonia oxidizers based on  
comparative 16S rRNA and amoA sequence analysis: implications for  
molecular diversity surveys  
Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)  
11092916  
JOURNAL 2 (bases 1 to 1526)  
Purthold,U., Pommerening-Roseer,A., Juretschko,S., Schmid,M.C.,  
Koops,H.P. and Wagner,M.  
Direct Submission  
Submitted (26-MAY-2000) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1..1526  
/organism="Nitrosomonas cryotolerans"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:44575"  
<1..>1526  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 90.7%; Score 1321.6; DB 1; Length 1526;  
Best Local Similarity 95.0%; Pred. No. 1e-13;  
Matches 1387; Conservative 0; Mismatches 69; Indels 4; Gaps 2;  
QY 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAATGCAAGCGACGACGCGATCTTGAT 60  
DB 21 ATTGAACGCTGGCGGCGATGCTTTACATGCAATGCAAGCGACGACGCGATCTTGAT 80  
QY 61 CTGGTGGCGATGGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 81 CTGGTGGCGATGGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 140  
QY 121 ACGCATGAAAGATGCTAATACCGCATATACCTTAAGAGGAAAGCAGGGGATCGAAA 180  
DB 141 ACGCATGAAAGATGCTAATACCGCATATACCTTAAGAGGAAAGCAGGGGATCGAAA 200  
QY 181 GACCTTGGCTTTTGAAGCGCGGATGCTGATTAAGTGAAGTGGGTTAAAGGCTTAC 240  
DB 201 GACCTTGGCTTTTGAAGCGCGGATGCTGATTAAGTGAAGTGGGTTAAAGGCTTAC 260  
QY 241 CAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 261 CAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 320  
QY 301 GCCCAGACTCTTAACGAGGCGACGATGAGGAAATTTTGAATGATGAGCGCAAGCTGATC 360  
DB 321 GCCCAGACTCTTAACGAGGCGACGATGAGGAAATTTTGAATGATGAGCGCAAGCTGATC 380  
QY 361 CAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGAAGGCTTTTCAATGAGAGAGA 420  
DB 381 CAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGAAGGCTTTTCAATGAGAGAGA 440  
QY 421 AAAGGTTACGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
DB 441 AAAGGTTACGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 500  
QY 481 TACGTGCCAGACGCCGCGTAAATGATGAGGTGCAAGCGTTAATCGAATTACTGGGCGT 540

DB 501 TACGTGCCAGACGCCGCGTAAATGATGAGGTGCAAGCGTTAATCGAATTACTGGGCGT 560  
QY 541 AAAGGTTACGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 561 AAAGGTTACGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 620  
QY 601 TCGGTTTGAATTAACAGGCTTGAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 621 TCGGTTTGAATTAACAGGCTTGAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 680  
QY 661 TGAATGCGTAGATGATGAGTGAAGAGCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
DB 681 TGAATGCGTAGATGATGAGTGAAGAGCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 740  
QY 721 GACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
DB 741 GACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 800  
QY 781 CTAAACGATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 801 CTAAACGATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 860  
QY 841 AGTTGACCGCTGGGAGGAGTACGATGAGTGAATTAAGGTTGATGATGATGATGATGATGAT 900  
DB 861 AGTTGACCGCTGGGAGGAGTACGATGAGTGAATTAAGGTTGATGATGATGATGATGATGAT 920  
QY 901 GCAACAACGCGTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 921 GCAACAACGCGTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 980  
QY 961 GACATGATGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017  
DB 981 GACATGATGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1040  
QY 1018 CTGGATGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077  
DB 1041 CTGGATGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1100  
QY 1078 ACCCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
DB 1101 ACCCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1160  
QY 1138 CCGGAGGAGGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
DB 1161 CCGGAGGAGGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1220  
QY 1198 TAATTAATGAGCGCGTAAAGAGGTTGCCAACCCGAGGAGGAGGAGGAGGAGGAGGAGG 1257  
DB 1221 TAATTAATGAGCGCGTAAAGAGGTTGCCAACCCGAGGAGGAGGAGGAGGAGGAGGAGG 1280  
QY 1258 GCGTGTAGTGCAGATGCGAGTCTGCAACTGCACTCCGTAAGTGGTAATTCGTAGTAAT 1317  
DB 1281 GCGTGTAGTGCAGATGCGAGTCTGCAACTGCACTCCGTAAGTGGTAATTCGTAGTAAT 1340  
QY 1318 CCGGATGACGATGCTGCGGTGAATACGTTCCGGGCTTTGTAACAACCGCCCGCTGACAC 1377  
DB 1341 CCGGATGACGATGCTGCGGTGAATACGTTCCGGGCTTTGTAACAACCGCCCGCTGACAC 1400  
QY 1378 CATGAGAGGAGGTTTCAACGAGAGCGATGATGATGATGATGATGATGATGATGATGATGAT 1437  
DB 1401 CATGAGAGGAGGTTTCAACGAGAGCGATGATGATGATGATGATGATGATGATGATGATGAT 1459  
QY 1438 TGAGATTCAATGAGGAGT 1457  
DB 1460 TGAGATTCAATGAGGAGT 1479

RESULT 31  
AF272414 1515 bp DNA linear BCT 05-DEC-2000  
LOCUS AF272414  
DEFINITION Nitrosomonas ureae 16S ribosomal RNA gene, partial sequence.  
ACCESSION AF272414



VERSION AF272414.1 GI:11545274  
KEYWORDS Nitrosomonas ureae  
SOURCE Nitrosomonas ureae  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1 (bases 1 to 1515)  
AUTHORS Purkhold, U., Pommerening-Roese, A., Juretschko, S., Schmidt, M.C., Koops, H.P. and Wagner, M.  
TITLE Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and amoA sequence analysis: implications for molecular diversity surveys  
JOURNAL Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)  
PUBMED 11057916  
REFERENCE 2 (bases 1 to 1515)  
AUTHORS Purkhold, U., Pommerening-Roese, A., Juretschko, S., Schmidt, M.C., Koops, H.P. and Wagner, M.  
TITLE Direct Substiation  
JOURNAL Submitted (26-MAY-2000) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1. .1515  
/organism="Nitrosomonas ureae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:44577"  
<1..>1515  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 90.6%; Score 1320.2; DB 1; Length 1515;  
Best Local Similarity 94.3%; Pred. No. 1..1e-13;  
Matches 1377; Conservative 0; Mismatches 80; Indels 3; Gaps 1;  
QY 1 ATTGAACGCTGGCGCGCATCTTTACACATGCAAGTGAACGCGACGCGATGCTTCAT 60  
DB 2 ATTCAAGCTGGCGCGCATCTTTACACATGCAAGTGAACGCGACGCGATGCTTCAT 61  
QY 61 CTGGTGGCGAGTGGCGGAGGCGGTGAGTATACATCGGAACGTCTTAAAGTGGGGATA 120  
DB 62 CTGGTGGCGAGTGGCGGAGGCGGTGAGTATACATCGGAACGTCTTAAAGTGGGGATA 121  
QY 121 ACGCATCGAAGATGCTATACCGCATATATCTTAAGAGAGAAAGCAGGGATCGA 180  
DB 122 ACGCATCGAAGATGCTATACCGCATATATCTTAAGAGAGAAAGTGGGGATCGA 181  
QY 181 GACCTTGGCGCTTTTGGAGCGCGCATGCTGATTAAGCTAGTGTGGGGTAAAGGCTTAC 240  
DB 182 GACCTTGGCGCTTTTGGAGCGCGCATGCTGATTAAGCTAGTGTGGGGTAAAGGCTTAC 241  
QY 241 CAAGCGACGATCATGATGTTGGTCTGAGAGACGACACCACTGGGACTGAGACACG 300  
DB 242 CAAGCGACGATCATGATGTTGGTCTGAGAGACGACACCACTGGGACTGAGACACG 301  
QY 301 GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATC 360  
DB 302 GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATC 361  
QY 361 CAGCAATGCGCGGTGAGTGAAGAGCGCTTGGGTTGTAAGCTTTCACTGAGAAAGA 420  
DB 362 CAGCAATGCGCGGTGAGTGAAGAGCGCTTGGGTTGTAAGCTTTCACTGAGAAAGA 421  
QY 421 AAAGGTACGGTAATTAATCGTACATGACGGTATCGACGAAGAAGCAACCGCTAAC 480  
DB 422 AAAGGTACGGTAATTAATCGTACATGACGGTATCGACGAAGAAGCAACCGCTAAC 481  
QY 481 TACGTGCCAGACCGCGGTAAATACGTAAGGTGCAAGCGTAAATCGAATTAATCTGGGCGT 540  
DB 482 TACGTGCCAGACCGCGGTAAATACGTAAGGTGCAAGCGTAAATCGAATTAATCTGGGCGT 541  
QY 541 AAAGGTGCGACGCGCGCTTTTGAATCATGATGTAATCCCGGCGCTTAACCTGGGAAT 600  
DB 542 AAAGGTGCGACGCGCGCTTTTGAATCATGATGTAATCCCGGCGCTTAACCTGGGAAT 601

QY 601 TGCGTTGAACCTACAGGCTAGAGTGGGAGGAGGAGTGAATTCATGTGAGCAG 660  
DB 602 TGCGTTGAACCTACAGGCTAGAGTGGGAGGAGGAGTGAATTCATGTGAGCAG 661  
QY 661 TGAATTCGCTAGATATAGGAAGAACATGATGGCGAAGGAGCGCTCTGGGTTAACT 720  
DB 662 TGAATTCGCTAGATATAGGAAGAACATGATGGCGAAGGAGCGCTCTGGGTTAACT 721  
QY 721 GACGCTATGACACGAACCGTGGGAGGAGAACAGATTAATCCCTGTAGTCCAGCC 780  
DB 722 GACGCTATGACACGAACCGTGGGAGGAGAACAGATTAATCCCTGTAGTCCAGCC 781  
QY 781 CTAAACGATGCTCACTAGTGTGGGCTTATTAAGCTTGGTAAAGACTAACCGTGA 840  
DB 782 CTAAACGATGCTCACTAGTGTGGGCTTATTAAGCTTGGTAAAGACTAACCGTGA 841  
QY 841 AGTTACCGCGCTGGGAGTACGCTGCAAGATTAACTCAAGGAATTGACGGGAGCC 900  
DB 842 AGTTACCGCGCTGGGAGTACGCTGCAAGATTAACTCAAGGAATTGACGGGAGCC 901  
QY 901 GCACAGCGGTGGATTAATGTGATTAATTCGATGCAACGCAAAACCTTACCTT 960  
DB 902 GCACAGCGGTGGATTAATGTGATTAATTCGATGCAACGCAAAACCTTACCTT 961  
QY 961 GACATGTAACGAATTTTCTAGAGATAGATAGTGTCT--TCGGGAACGCTAACAGGTG 1017  
DB 962 GACATGTAAGAAATATTTCAAGATGATTTGTGTCTCAAAAGAAATCTAAACAGGTG 1021  
QY 1018 CTGCATGCGCTGCTGACGCTGCTGCTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1077  
DB 1022 CTGCATGCGCTGCTGACGCTGCTGCTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1081  
QY 1078 ACCCTTGCATTAATATGCAATCATTTGGTGGGCACTTAATGAGACTGGCGGAGCAA 1137  
DB 1082 ACCCTTGCATTAATATGCAATCATTTAGTTGGGCACTTAATGAGACTGGCGGAGTAA 1141  
QY 1138 CCGAGAGAAAGTGGGAGTGAAGTCAAGTCTCATGACCTTATGGGTAGGGCTTACACG 1197  
DB 1142 CCGAGAGAAAGTGGGAGTGAAGTCAAGTCTCATGACCTTATGGGTAGGGCTTACACG 1201  
QY 1198 TAAATCAATGCGCGCTACAGAGGCTTGCACACCGCGAGGGGAGCTAATCTCAGAAAGC 1257  
DB 1202 TAAATCAATGCGCGCTACAGAGGCTTGCACACCGCGAGGGGAGCTAATCTCAGAAAGC 1261  
QY 1258 GCGTGTGATGTCGGATCGGAGTCTGCACTGCACTCCGTGAAGTGGGAATGGCTAGTAT 1317  
DB 1262 GCGTGTGATGTCGGATCGGAGTCTGCACTGCACTCCGTGAAGTGGGAATGGCTAGTAT 1321  
QY 1318 GCGGATGAGCATGTCGCGGTGAATACGTTCCGGGCTTTGTAACACACCGCCGTACAC 1377  
DB 1322 GCGGATGAGCATGTCGCGGTGAATACGTTCCGGGCTTTGTAACACACCGCCGTACAC 1381  
QY 1378 CATGGAGTGGGTTTACACAGAGGAGTACTTAACCGTAAGAGAGGCGCTTGCACGG 1437  
DB 1382 CATGGAGTGGGTTTACACAGAGGAGTACTTAACCGTAANNNGGCGCTTGCACGG 1441  
QY 1438 TGAATTTATGACTGGGGTG 1457  
DB 1442 TGAATTTATGACTGGGGTG 1461  
RESULT 32  
AJ621026 1454 bp DNA linear BCT 08-JUN-2004  
DEFINITION Nitrosomonas sp. I879A3 16S rRNA gene, isolate I879A3.  
ACCESSION AJ621026  
VERSION AJ621026.1 GI:4094845  
KEYWORDS 16S ribosomal RNA; 16S rRNA gene.  
SOURCE Nitrosomonas sp. I879A3  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1



**AUTHORS** Nielsen, S., Baer Gijlsen, M.J., Revsbech, N., Laandroek, H.J. and Bollmann, A.  
**TITLE** Influence of heterotrophic bacteria on the activity and growth behavior of ammonia-oxidizing bacteria  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 1454)  
**AUTHORS** Bollmann, A.  
**JOURNAL** Direct Submission  
 Submitted (15-JAN-2004) Bollmann A., Department for Microbial Ecology, NIOO-KNAW Centre for Limnology, Rijksstraatweg 6, 3631 AC Nieuwerstadiu, NETHERLANDS  
**FEATURES** Location/Qualifiers  
 source  
 1. 1454  
 /organism="Nitrosomonas sp. 1s79A3"  
 /mol\_type="genomic DNA"  
 /isolate="1s79A3"  
 /isolation\_source="fresh water sediment"  
 /db\_xref="taxon:261292"  
 /country="Netherlands; Lake Drontermeer"  
 1. 1449  
 /gene="16S rRNA"  
 1. 1449  
 /gene="16S rRNA"  
 /product="16S ribosomal RNA"  
**ORIGIN**  
 Query Match 90.3%; Score 1315.4; DB 1; Length 1454;  
 Best Local Similarity 95.5%; Pred. No. 1.3e-13;  
 Matches 1387; Conservative 0; Mismatches 61; Indels 5; Gaps 3;

1 ATTGAACGCTGGCGGCGATCTTTACACATGCAAGTGAACGGGAGACGGAGCTTTCAT 60  
 4 ATTGAACGCTGGCGGCGATCTTTACACATGCAAGTGAACGGGAGACGGAGCTTTCAT 63  
 61 CTGGTGGCGAGTGGCGGAGCGGTGATGATGATCGGAACGTATCCAGAAAGAGGGGGT 120  
 64 CTGGTGGCGAGTGGCGGAGCGGTGATGATGATCGGAACGTATCCAGAAAGAGGGGGT 122  
 121 AGCGATCGAAAGATGCTAATACCGCATATCTTAAGAGGAAAGAGGGAGTGA 180  
 123 AGCGATCGAAAGATGCTAATACCGCATATCTTAAGAGGAAAGAGGGAGTGA 182  
 181 GACCTTGGCTTTTGGAGCGCGCGATGCTGATTAAGCTTGGTAAAGGCTTAC 240  
 183 GACCTTATGCTTTTGGAGCGCGCGATGCTGATTAAGCTTGGTAAAGGCTTAC 242  
 241 CAAGCGAGCATAGTATGTTGTTCTGAGAGAGACGACGACACTGGGACTGAGACAG 300  
 243 CAAGCGAGCATAGTATGTTGTTCTGAGAGAGCGGCGACGACACTGGGACTGAGACAG 302  
 301 GCCCACTCTCTACGGGAGGACGACAGTGGGAAATTTTGAACAATGGCGGACTGATC 360  
 303 GCCCACTCTCTACGGGAGGACGACAGTGGGAAATTTTGAACAATGGCGGACTGATC 362  
 361 CAGCAATGCCGCTGAGTGAAGAAAGGCTTCGGGTGTAAAGCTTTTCAGTCGAGAGA 420  
 363 CAGCAATGCCGCTGAGTGAAGAAAGGCTTCGGGTGTAAAGCTTTTCAGTCGAGAGA 422  
 421 AAAGGTACGGTAATATCTGACTATGACGATTCGACAGAAAGAACCGGCTTAC 480  
 423 AAAGGTGTGATCAATATCACTATACGATATCGACAGAAAGAACCGGCTTAC 482  
 481 TAGGTGCACAGCGCGGCTAATACGTAAGGCTGCAACGTTAATCGAATTTACTGGGCT 540  
 483 TAGGTGCACAGCGCGGCTAATACGTAAGGCTGCAACGTTAATCGAATTTACTGGGCT 542  
 541 AAAGGTGCCAGCGGCTTTGTAAAGTCAAGTGAATATCCCGGGCTTAACTGGGAAT 600  
 543 AAAGGTGCCAGCGGCTTTGTAAAGTCAAGTGAATATCCCGGGCTTAACTGGGAAT 602  
 601 TGGCTTTGAAACTCAAGGCTAGTGTGGCAGAGGAGGTGAATTCATGTCTACAG 660  
 603 TGGCTTTGAAACTCAAGGCTAGTGTGGCAGAGGAGGTGAATTCATGTCTACAG 662

661 TGAATGCTGATAGATATGGAAGAACATCGATGGCGAAGGAGCGCTCCGGTTAACT 720  
 663 TGAATGCTGATAGATATGGAAGAACATCGATGGCGAAGGAGCGCTCCGGTTAACT 722  
 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGGTATCCAGCGC 780  
 723 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGGTATCCAGCGC 782  
 781 CTAAACGATGCTAATAGTTGTTGGCCTTATTAGCCTTGGTATACGAAGCTTAACCGCT 840  
 783 CTAAACGATGCTAATAGTTGTTGGCCTTATTAGCCTTGGTATACGAAGCTTAACCGCT 842  
 841 AGTTAGCCGCTGGGAGATACCGTGGCAAGATTAACTCAAGGAATTTGAAGGGAGACC 900  
 843 AGTTAGCCGCTGGGAGATACCGTGGCAAGATTAACTCAAGGAATTTGAAGGGAGACC 902  
 901 GCAAGAGCGGTGATTAATGATGATTAATTCGATGCAACGCAAAAACCTTACCTTACCCT 960  
 903 GCAAGAGCGGTGATTAATGATGATTAATTCGATGCAACGCAAAAACCTTACCTTACCCT 962  
 961 GACATGTAGCGAATTTTCTAGAGATGATTAAGTCT--TCGGAAACGCTTAACACAGT 1017  
 963 GACATGTAGCGAATTTTCTAGAGATGATGATTAAGTCTCGAAAGAGATTTGAACACAGT 1022  
 1018 CTGCATGCTGCTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
 1023 CTGCATGCTGCTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082  
 1078 ACCCTTGTCAATTAATGCTCAATTTGTTGGGCACTTAAATGACACTCCGGTGA 1137  
 1083 ACCCTTGTCAATTAATGCTCAATTTGTTGGGCACTTAAATGACACTCCGGTGA 1142  
 1138 CCGAGAGAGTGGGATGACGTCATGCTCATGCGCTTATGGGTAGGCTTCAACG 1197  
 1143 CCGAGAGAGTGGGATGACGTCATGCTCATGCGCTTATGGGTAGGCTTCAACG 1202  
 1198 TAATCAATGCGCGCTTACAGAGGTTGCCAACCCCGAGAGGGAGCTTAATTCAGAAAG 1257  
 1203 TAATCAATGCGCGCTTACAGAGGTTGCCAACCCCGAGAGGGAGCTTAATTCAGAAAG 1262  
 1258 GCGGTGATCCGGATCGAGATCTGCACTGCACTCGTGAAGTCCGAATCGCTAAT 1317  
 1263 GCGGTGATCCGGATCGAGATCTGCACTGCACTCGTGAAGTCCGAATCGCTAAT 1322  
 1318 CCGGATCAGACATGTCGCGGTGAATCGTTCCGGGCTTTGTAACACACCGCCGTCACAC 1377  
 1323 CCGGATCAGACATGTCGCGGTGAATCGTTCCGGGCTTTGTAACACACCGCCGTCACAC 1382  
 1378 CATGGAGTGGGTTTCAACAGAGCGATGATTAACCGTAAGAGAGGGCTTGGCCACG 1437  
 1383 CATGGAGTGGGTTTCAACAGAGCGATGATTAACCGTAAGAGAGGGCTTGGCCACG 1441  
 1438 TGAATTCATGAC 1450  
 1442 TGAATTCATGAC 1454

**RESULT 33**  
 AY123794 1501 bp DNA linear BCT 23-SBP-2003  
 LOCUS Nitrosomonas sp. Nm143 16S ribosomal RNA gene, partial sequence.  
 DEFINITION AY123794  
 ACCESSION AY123794  
 VERSION AY123794.1 GI:24474423  
 KEYWORDS  
 SOURCE Nitrosomonas sp. Nm143  
 ORGANISM Nitrosomonas sp. Nm143  
 Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
 Nitrosomonadaceae; Nitrosomonas.  
 REFERENCE 1 (bases 1 to 1501)  
 Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Rosier, A. and Koepe, H. P.  
 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial

ammonia-oxidizing isolates: extension of the dataset and proposal of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)

JOURNAL  
PUBMED  
13130037  
2 (bases 1 to 1501)  
Purkhold, U., Timmermann, G., Koope, H.-P., Pommerening-Roeser, A. and Wagner, M.

TITLE  
Journal  
Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany

FEATURES  
source  
1. 1501  
/organism="Nitrosomonas sp. Nm143"  
/mol\_type="genomic DNA"  
/strain="Nm143"  
/db\_xref="taxon:200123"  
<1. .>1501  
/product="16S ribosomal RNA"

ORIGIN  
rRNA

Query Match 90.2%; Score 1314.8; DB 1; Length 1501;  
Best Local Similarly 95.1%; Pred. No. 1.3e-13;  
Matches 1390; Conservative 0; Mismatches 67; Indels 5; Gaps 3;

1 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGACGCGACGATGCTTGCAT 60  
3 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGACGCGACGATGCTTGCAT 62  
61 CTGTGGCGAGTGGCGGACGCGGTGAGTAATGCAATCGAACTATCGAAGAGGGGGTA 120  
63 CTGGTGGCGAGTGGCGGACGCGGTGAGTAATGCAATCGAACTATCGAAGAGGGGGTA 122  
121 AGCGATCGAAGATGCTTAATCCGCAATTAATCTTAAGAGAGAAAGCGGGATCGAAA 180  
123 AGCGATCGAAGATGCTTAATCCGCAATTAATCTTAAGAGAGAAAGCGGGATCGAAA 182  
181 GACCTTGGCGCTTTTGAAGCGCGCGATGCTGATTAAGCTAGTGGGGTAAAGCCCTAC 240  
183 GACCTTGGCGCTTTTGAAGCGCGCGATGCTGATTAAGCTAGTGGGGTAAAGCCCTAC 242  
241 CAAGCGACGATCAGTAAGTGTGCTGAGAGACGACGACGACCACTGGGACTGAGACAG 300  
243 CAAGCGACGATCAGTAAGTGTGCTGAGAGACGACGACGACCACTGGGACTGAGACAG 302  
301 GCCGACGATCTCTACGCGAGGACGACGATGGGGAAATTTGACAAATGGGCGCAAGCTGATC 360  
303 GCCGACGATCTCTACGCGAGGACGACGATGGGGAAATTTGACAAATGGGCGCAAGCTGATC 362  
361 CAAGATGCGCGGTGAGTAAGAGGCGCTTGGGGTGTAAAGCTCTTTCAGTCGAGAAAGA 420  
363 CAAGATGCGCGGTGAGTAAGAGGCGCTTGGGGTGTAAAGCTCTTTCAGTCGAGAAAGA 422  
421 AAAGGTTACGTTAAATCGTACTCATGACGCGATTCGACAGAAAGACCGGCTAAC 480  
423 AAAGGTTACGTTAAATCGTACTCATGACGCGATTCGACAGAAAGACCGGCTAAC 482  
481 TACGTCGACGACGCGCGGTGAGTAAGAGGCGCTTTCAGTCGAGAAAGA 540  
483 TACGTCGACGACGCGCGGTGAGTAAGAGGCGCTTTCAGTCGAGAAAGA 542  
541 AAAGGTTACGTTAAATCGTACTCATGAGTGAATCCCGGGGCTTAACCTGGGAAT 600  
543 AAAGGTTACGTTAAATCGTACTCATGAGTGAATCCCGGGGCTTAACCTGGGAAT 602  
601 TCGGTTTGAATCAAGGCTTAAGTGTGCGACAGAGAGTGAATTCATGCTGAGCAG 660  
603 TCGGTTTGAATCAAGGCTTAAGTGTGCGACAGAGAGTGAATTCATGCTGAGCAG 662  
661 TGAATTCGCTAGATGATGAGAAACATGATGGCGAAGGCGCTCTCGGCTTAACCT 720  
663 TGAATTCGCTAGATGATGAGAAACATGATGGCGAAGGCGCTCTCGGCTTAACCT 722  
721 GACGCTCATGACGAAGGCTGGGAGCAAAACAGATTAGATACCTGTAGTCCAGCC 780

|||||  
723 GACGCTCAGGACGAAACGTTGGGACGAAACAGGATTAAGTACCCTGTAGTCCAGCC 782  
781 CTAAACGATGTCATTAAGTGTGGGCTTTATGAGCTTGGTGAACGAACTAACCGCTGA 840  
783 CTAAACGATGTCATTAAGTGTGGGCTTTATGAGCTTGGTGAACGAACTAACCGCTGA 842  
841 AGTTGACGCGCTGGGGAATGAGTGGCAAGATTAATACTCAAGAGATTTGACGGGAGCC 900  
843 AGTTGACGCGCTGGGGAATGAGTGGCAAGATTAATACTCAAGAGATTTGACGGGAGCC 902  
901 GCACAGGCGGTGATTAATGAGTGAATTAATTCGATGCAACGCAAAAACCTTAACCTT 960  
903 GCACAGGCGGTGATTAATGAGTGAATTAATTCGATGCAACGCAAAAACCTTAACCTT 962  
961 GACATGT-AGCGAATTTTCTAGAGATGATTAAGTCT--TTCCGGAAGCTTAACAGAGT 1016  
963 GACATGTACCGGAAGCTTAAGTGAAGTGAAGTGGCGCCGAAAGGAGCGGTAAACAGAGT 1022  
1017 GCTGATGCGTGTGTCAGCTCGTGTGTCGATGATGTTAGTCCGCAACGAGGC 1076  
1023 GCTGATGCGTGTGTCAGCTCGTGTGTCGATGATGTTAGTCCGCAACGAGGC 1082  
1077 AACCTTGTCAATTAATGTCATCAATTTGTTGGGCACTTAATGAGACTGCGGTGACAA 1136  
1083 AACCTTGTCAATTAATGTCATCAATTTAGTTGGGCACTTAATGAGACTGCGGTGACAA 1142  
1137 ACCGAGGAGAGTGGGAGATGACGT-CAAGTCTTCAATGCGCTTATGAGTGGCTTCA 1195  
1143 ACCGAGGAGAGTGGGAGATGACGTCAAGTCTTCAATGCGCTTATGAGTGGCTTCA 1202  
1196 GGTAAATCAATGAGGCGGTGACAGAGGTTGCCAACCCGCGAGGAGGAGTAAATTCAGAAA 1255  
1203 GGTAAATCAATGAGGCGGTGACAGAGGTTGCCAACCCGCGAGGAGGAGTAAATTCAGAAA 1262  
1256 GCGGCTGTAGTCCGATCGGAGTGTGCAACTGCACTCGTGAATGCGAATTCGTAATA 1315  
1263 GCGGCTGTAGTCCGATCGGAGTGTGCAACTGCACTCGTGAATGCGAATTCGTAATA 1322  
1316 ATCCGATGATGATGATGTCGCGGTGAATTCGTTCCCGGCTCTTGTACACACCGCCGCTAC 1375  
1323 ATCCGATGATGATGATGTCGCGGTGAATTCGTTCCCGGCTCTTGTACACACCGCCGCTAC 1382  
1376 ACCATGGAGAGTGGGTTTACCAAGAGAGTGTACTAATCGTAAGAGAGGCGCTTGCAC 1435  
1383 ACCATGGAGAGTGGGTTTACCAAGAGAGTGTACTAATCGTAAGAGAGGCGCTTGCAC 1442  
1436 GGTGAGATTCAATGACTGGGGTG 1457  
1443 GGTGAGATTCAATGACTGGGGTG 1464  
RESULT 34  
AY123806 1497 bp DNA linear BCT 23-SEP-2003  
LOCUS  
DEFINITION Nitrosospira sp. Kaz 16S ribosomal RNA gene, partial sequence.  
ACCESSION AY123806  
VERSION AY123806.1 GI:24474435  
KEYWORDS  
SOURCE  
ORGANISM Nitrosospira sp. Kaz  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira.  
REFERENCE  
1 (bases 1 to 1497)  
Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roeser, A. and Koope, H.-P.  
TITLE  
Journal  
Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
PUBMED  
13130037  
2 (bases 1 to 1497)  
Purkhold, U., Timmermann, G., Koope, H.-P., Pommerening-Roeser, A. and



ORIGIN

rRNA

/organism="Nitrosospirra sp. Nsp57"  
/mol\_type="genomic DNA"  
/strain="Nsp57"  
/db\_xref="taxon:200118"  
<1..>1497  
/product="16S ribosomal RNA"

Query Match 89.7%; Score 1306.8; DB 1; Length 1497;  
Best Local Similarity 94.4%; Pred. No. 1.8e-13;  
Matches 1378; Conservative 0; Mismatches 77; Indels 5; Gaps 2;

1 ATTGAACGCTGGCGGATCTTTACATGCAAGTCGAACGGCAGCAGCATGCTTCAT 60  
3 ATTGAACGCTGGCGGATCTTTACATGCAAGTCGAACGGCAGCAGCATGCTTCAT 60  
61 CTGTGGCGAGTGGCGGAGCGGTGAGTAAATGCAATCGGAAAGTATCCAGAAAGGCGGAT 120  
61 CTGTGGCGAGTGGCGGAGCGGTGAGTAAATGCAATCGGAAAGTATCCAGAAAGGCGGAT 120  
121 AGCGATCGAAAGATGTGCTTAATCCGCAATTAATCTTAAGAGAGAAAGCGGGATCGAA 180  
121 AGCGATCGAAAGATGTGCTTAATCCGCAATTAATCTTAAGAGAGAAAGCGGGATCGAA 180  
181 GACCTTGGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTAAAGGCTTAC 240  
181 GACCTTGGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTAAAGGCTTAC 240  
241 CAAGCGACGATCAGTAGTGTGCTGAGAGAGACGACGACCACTGGGACTGAGACAGC 300  
241 CAAGCGACGATCAGTAGTGTGCTGAGAGAGACGACGACCACTGGGACTGAGACAGC 300  
301 GCCCAGACTCTTCAACGGAGGAGCAGATGGGGAATTTTGAACAATGGGGAATCCCTGATC 360  
301 GCCCAGACTCTTCAACGGAGGAGCAGATGGGGAATTTTGAACAATGGGGAATCCCTGATC 360  
361 CAGCATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCACTGAGAGAGA 420  
361 CAGCATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCACTGAGAGAGA 420  
421 AAAGGTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 480  
421 AAAGGTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 480  
481 AACGCTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 480  
481 AACGCTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 480  
481 TACGCTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 540  
481 TACGCTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 540  
541 AAAAGGTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 600  
541 AAAAGGTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 600  
541 AAAGGTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 600  
541 AAAGGTTACGCTAAATTAATCTGATCATGACGCTATCGACAGAAAGACACCGGCTAAC 600  
601 TGGGTTTGAACCTACAGAGCTAGAGTGGGAGAGGAGGTGAAATTCATGTGTACG 660  
601 TGGGTTTGAACCTACAGAGCTAGAGTGGGAGAGGAGGTGAAATTCATGTGTACG 660  
661 TGGGTTTGAACCTACAGAGCTAGAGTGGGAGAGGAGGTGAAATTCATGTGTACG 660  
661 TGGGTTTGAACCTACAGAGCTAGAGTGGGAGAGGAGGTGAAATTCATGTGTACG 660  
721 GACGCTCATGACAGAAAGGCTGGGAGAGCAACAGATTAATCCCTGGTATGCAAGCC 780  
721 GACGCTCATGACAGAAAGGCTGGGAGAGCAACAGATTAATCCCTGGTATGCAAGCC 780  
781 GACGCTCATGACAGAAAGGCTGGGAGAGCAACAGATTAATCCCTGGTATGCAAGCC 780  
781 GACGCTCATGACAGAAAGGCTGGGAGAGCAACAGATTAATCCCTGGTATGCAAGCC 780  
841 AGTTACCGCTTGGGAGAGTACGCTGCAAGATTTAAATCTCAAGAGATTTGACGGGAGCC 900  
841 AGTTACCGCTTGGGAGAGTACGCTGCAAGATTTAAATCTCAAGAGATTTGACGGGAGCC 900  
901 GCACAAAGCGGTGATTAATGATTAATTCATGACAGCAAGCAAACTTACCTACCTT 960

Db

901 GCACAAAGCGGTGATTAATGATTAATTCATGACAGCAAGCAAACTTACCTACCTT 960

Qy

961 GACATGTACCGAATTTTCTAGAGATAGATTAGTGC---TTGGGAAACGCTAACACAGGTG 1017

Db

961 GACATGTACCGAATTTTCTAGAGATAGATTAGTGC---TTGGGAAACGCTAACACAGGTG 1020

Qy

1018 CTGCATGAGCTGCGTACAGTGTGCTGAGATGTTGGGATTAAGTCCGCAACAGAGCCCA 1077

Db

1021 CTGCATGAGCTGCGTACAGTGTGCTGAGATGTTGGGATTAAGTCCGCAACAGAGCCCA 1080

Qy

1078 ACCCTTGTCAATTAATTCGATCATCATTTGGTGGGCACTTAATGAGACTGCGGTGACAA 1137

Db

1081 ACCCTTGTCAATTAATTCGATCATCATTTAGTGGGCACTTAATGAAATGCGCGGTGACAA 1140

Qy

1138 CCGAGAGAAAGTGGGAGATGAGCTCAAGTCTCATAGTCCCTTATGGGTAGGCTTACACG 1197

Db

1141 CCGAGAGAAAGTGGGAGATGAGCTCAAGTCTCATAGTCCCTTATGGGTAGGCTTACACG 1200

Qy

1198 TAAATCAATGGCGGTACAGAGGCTTGGCAACCCGCGAGAGGAGACTAATCTCAGAAAGC 1257

Db

1201 TAAATCAATGGCGGTACAGAGGCTTGGCAACCCGCGAGAGGAGACTAATCTCAGAAAGC 1260

Qy

1258 GCGTGTAGTCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 1317

Db

1261 GCGTGTAGTCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 1320

Qy

1318 CCGGATACGATGTCGCGGTGAAATACCTTCCGGGTCTTATACACACGCGCGGTACAC 1377

Db

1321 CCGGATACGATGTCGCGGTGAAATACCTTCCGGGTCTTATACACACGCGCGGTACAC 1380

Qy

1378 CATGGAGTGGGTTTCAACGAGAGAGTGTAAACCGTAAGAGAGGCGCTTGGCAGCG 1437

Db

1381 CATGGAGTGGGTTTCAACGAGAGAGTGTAAACCGTAAGAGAGGCGCTTGGCAGCG 1440

Qy

1438 TGAGATTCAATGACTGGGGGTG 1457

Db

1441 TGAGATTCAATGACTGGGGGTG 1460

RESULT 36

AY123789 1497 bp DNA linear BCT 23-SEP-2003

LOCUS Nitrosospirra sp. Nsp58 16S ribosomal RNA gene, partial sequence.

DEFINITION AY123789

ACCESSION AY123789.1 GI:24474418

VERSION

KEYWORDS

SOURCE

ORGANISM Nitrosospirra sp. Nsp58

REFERENCE 1 (bases 1 to 1497) Koops,H.P., Timmermann,G., Pommerening-Roseer,A. and Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosospirra.

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 1497

/organism="Nitrosospirra sp. Nsp58"

/mol\_type="genomic DNA"

/strain="Nsp58"

/db\_xref="taxon:200116"

<1..>1497

/product="16S ribosomal RNA"

## ORIGIN

Query Match 89.6%; Score 1305.2; DB 1; Length 1497;

Best Local Similarity 94.3%; Pred. No. 1.9e-13;

Matches 1377; Conservative 0; Mismatches 78; Indels 5; Gaps 2;

QY 1 ATTGAACGCTGGCGGAGTCTTTTACATGCAAGTCGAAACGGACGACGAGATCTTGCAAT 60  
 DB 3 ATTGAACGCTGGCGGAGTCTTTTACATGCAAGTCGAAACGGACGACGAGATCTTGCAAT 60  
 QY 61 CTGGTGGCCGAGTGGCCGACCGGCTGATTAATGCAATCGGAACGTATCCAGAAAGAGGGGGAT 120  
 DB 61 CTGGTGGCCGAGTGGCCGACCGGCTGATTAATGCAATCGGAACGTATCCAGAAAGAGGGGGAT 120  
 QY 121 AGGCATCGAAAGATGTGCTAATACCGCATATCTTAAGGAGGAAAGACAGGGGATGAAA 180  
 DB 121 AGGCATCGAAAGATGTGCTAATACCGCATATCTTAAGGAGGAAAGACAGGGGATGAAA 180  
 QY 181 GACCTTGCCTTTTGAAGCGGCGCATGTCTGATTAAGTGTGGGTAAAGGCTTAC 240  
 DB 181 GACCTTGCCTTTTGAAGCGGCGCATGTCTGATTAAGTGTGGGTAAAGGCTTAC 240  
 QY 241 CAAGGCGACATCAATGTTGCTTGAAGAGACGACACGACCACTGGGACCTGAGACAG 300  
 DB 241 CAAGGCGACATCAATGTTGCTTGAAGAGACGACACGACCACTGGGACCTGAGACAG 300  
 QY 301 GCCCAGACTCTTAAGGAGGACGACGAGGGGAAATTTTGAACAATGGGCGCAAGCTGATC 360  
 DB 301 GCCCAGACTCTTAAGGAGGACGACGAGGGGAAATTTTGAACAATGGGCGGAAACCTGATC 360  
 QY 361 CAGCAATGCCGCGTGAATGAAGAAAGGCTTCGGGTGTAAAGCTCTTTCAGTGAAGAA 420  
 DB 361 CAGCAATGCCGCGTGAATGAAGAAAGGCTTCGGGTGTAAAGCTCTTTCAGCAGGAAAGA 420  
 QY 421 AAAGGTTACGCTAATATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 480  
 DB 421 AAAGGTTACGCTAATATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 480  
 QY 481 TAGGTGCACGACGCGCGGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 540  
 DB 481 TAGGTGCACGACGCGCGGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 540  
 QY 541 AAAGGTCGACGACGCGCTTTGTAAGTCAATGTGAATCCCGGCGCTTAACCTGGGAAT 600  
 DB 541 AAAGGTCGACGACGCGCTTTGTAAGTCAATGTGAATCCCGGCGCTTAACCTGGGAAT 600  
 QY 601 TGGCTTTGAAACTCAAGGCTAAGTGTGCAAGGAGGAGTGGAAATTCATGTGTACAG 660  
 DB 601 TGGCTTTGAAACTCAAGGCTAAGTGTGCAAGGAGGAGTGGAAATTCATGTGTACAG 660  
 QY 661 TGAATATCGTAAAGATGTGAAGAAACATGATGAGCAAGGACGCTCGGAGTTAACT 720  
 DB 661 TGAATATCGTAAAGATGTGAAGAAACATGATGAGCAAGGACGCTCGGAGTTAACT 720  
 QY 721 GAGCTCATGACGAAACCGTGGGAGCAAAACAGATTAATGATCCTGTGTACAGCC 780  
 DB 721 GAGCTCATGACGAAACCGTGGGAGCAAAACAGATTAATGATCCTGTGTACAGCC 780  
 QY 781 CTAAACGATGCACTAGTTGTGGGCTTATTAGGCTTGTATACGAAGCTTAAGCGCTGA 840  
 DB 781 CTAAACGATGCACTAGTTGTGGGCTTATTAGGCTTGTATACGAAGCTTAAGCGCTGA 840  
 QY 841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAAACTCAAGGAATGACGGGAGCC 900  
 DB 841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAAACTCAAGGAATGACGGGAGCC 900  
 QY 901 GCACAAAGCGGTGATTAATGATTAATTCGATGCAACGCGAAGAAACCTTAACCTT 960  
 DB 901 GCACAAAGCGGTGATTAATGATTAATTCGATGCAACGCGAAGAAACCTTAACCTT 960  
 QY 961 GACATGTAGCGAATTTTCTAGAGATAGTTAGTC---TTGGGAGCGCTTAACAGAGTG 1017  
 DB 961 GACATGTAGCGAATTTTCTAGAGATAGTTAGTC---TTGGGAGCGCTTAACAGAGTG 1017  
 QY 1017 GACATGTAGCGAATTTTCTAGAGATAGTTAGTC---TTGGGAGCGCTTAACAGAGTG 1020  
 DB 1017 GACATGTAGCGAATTTTCTAGAGATAGTTAGTC---TTGGGAGCGCTTAACAGAGTG 1020

QY 1018 CTGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077  
 DB 1021 CTGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
 QY 1078 ACCCTGTCAATTAATGCTCAATTAATGCTCAATTAATGCTCAATTAATGCTCAATTAATG 1137  
 DB 1081 ACCCTGTCAATTAATGCTCAATTAATGCTCAATTAATGCTCAATTAATGCTCAATTAATG 1140  
 QY 1138 CCGAGGAAAGTGGGAGTACGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATG 1197  
 DB 1141 CCGAGGAAAGTGGGAGTACGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATG 1200  
 QY 1198 TAATCAATGCGGCTGCAAGAGGCTTCCAAACCGGAGGGGAGCTAATCTCAGAAAGC 1257  
 DB 1201 TAATCAATGCGGCTGCAAGAGGCTTCCAAACCGGAGGGGAGCTAATCTCAGAAAGC 1260  
 QY 1258 GGGTGTAGTCCGAGTCGAGTCTGCAATCTGCAATCTGCAATCTGCAATCTGCAATCTGCAAT 1317  
 DB 1261 GGGTGTAGTCCGAGTCGAGTCTGCAATCTGCAATCTGCAATCTGCAATCTGCAATCTGCAAT 1320  
 QY 1318 CCGGATCAGCATGTCGCGGTGAATGCTTCCGAGTCTTGTACACACCGCCGCTACAC 1377  
 DB 1321 CCGGATCAGCATGTCGCGGTGAATGCTTCCGAGTCTTGTACACACCGCCGCTACAC 1380  
 QY 1378 CATGGAGTGGGTTTACCAAGAGAGTGAATCCGTTAAGAGAGGGGCTTGCCACCG 1437  
 DB 1381 CATGGAGTGGGTTTACCAAGAGAGTGAATCCGTTAAGAGAGGGGCTTGCCACCG 1440  
 QY 1438 TGAATTCATGACTGGGGTG 1457  
 DB 1441 TGAATTCATGACTGGGGTG 1460

## RESULT 37

AY123793

LOCUS

AY123793

DEFINITION

AY123793

ACCESSION

AY123793.1

VERSION

GI:24474422

KEYWORDS

NITROSPIRA SP. NPS

SOURCE

NITROSPIRA SP. NPS

ORGANISM

NITROSPIRA SP. NPS

REFERENCE

Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roesser, A. and

Koops, H. P.

16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial

ammonia-oxidizing isolates: extension of the dataset and proposal

of a new lineage within the Nitrospira

Inc. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)

JOURNAL

PUBMED

13130037

AUTHORS

Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roesser, A. and

Wagner, M.

Direct Submission

Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische

Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany

Location/Qualifiers

1. 1497

/organism="Nitrospira sp. Nps"

/mol\_type="genomic DNA"

/strain="Nps5"

/db\_xref="taxon:200119"

&lt;1. &gt;1497

/product="16S ribosomal RNA"

## ORIGIN

Query Match 89.6%; Score 1305.2; DB 1; Length 1497;  
 Best Local Similarity 94.3%; Pred. No. 1.9e-13;  
 Matches 1377; Conservative 0; Mismatches 78; Indels 5; Gaps 2;

```
QY 1 ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAACGGCAGACCGATGCTTGCAT 60
DB 3 ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAACGGCAGACCGGCGCAAC-- 60
QY 61 CTGTGGCGAGTGGCGGACGGGTGAGTAAATGCAATCGAACCTATCCAGAAAGAGGGGGTGA 120
DB 61 CTGTGGCGAGTGGCGGACGGGTGAGTAAATGCAATCGAACCTATCCAGAAAGAGGGGGTGA 120
QY 121 AGCGATCGAAAGATGTGCTTAATACCGCATATATCTCTAAGAGAGAAAGCGGGGATCGAAA 180
DB 121 AGCGATCGAAAGATGTGCTTAATACCGCATATATCTCTAAGAGAGAAAGCGGGGATCGCAA 180
QY 181 GACCTTGGCCTTTGGAGCGCGCATGTCTGATTAGCTAGTGGTGGGTAAAGGCTTAC 240
DB 181 GACCTTGGCCTTTGGAGCGCGCATGTCTGATTAGCTAGTGGTGGGTAAAGGCTTAC 240
QY 241 CAAGCGACGATCAGTATGTTGGTCTGAGAGAGACGACACGACCACTGGGACTGAGACACG 300
DB 241 CAAGCGCTTCGATCAGTATGTTGGTCTGAGAGAGACGACACGACCACTGGGACTGAGACACG 300
QY 301 GCCCAGACTCTTACGGGAGGCGACGATGGGGAAATTTTGGACATATGGGCGCAACCTGATC 360
DB 301 GCCCAGACTCTTACGGGAGGCGACGATGGGGAAATTTTGGACATATGGGCGCAACCTGATC 360
QY 361 CAGCAATGCGGCGGTGAGTGAAGAGGCGCTTGGGCTTGAAGCTCTTTCAGTCGAGAAAG 420
DB 361 CAGCAATGCGGCGGTGAGTGAAGAGGCGCTTGGGCTTGAAGCTCTTTCAGTCGAGAAAG 420
QY 421 AAAGGTTACGTTAAATATCTGACTCATGACGCTATCGACAGAAAGAACCGGCTAAC 480
DB 421 AAAGGTTACGTTAAATATCTGACTCATGACGCTATCGACAGAAAGAACCGGCTAAC 480
QY 481 TACGTTGCCAGCAGCCCGGCTAATACGTAGGGTGCAGGCTTAAATCGAAATTAATCTGGGCT 540
DB 481 TACGTTGCCAGCAGCCCGGCTAATACGTAGGGTGCAGGCTTAAATCGAAATTAATCTGGGCT 540
QY 541 AAAGGGTGGCAGGCGGCTTTGTAAGTCAATGATGTGAATCCCGCGGCTTAACTGGGAAAT 600
DB 541 AAAGGGTGGCAGGCGGCTTTGTAAGTCAATGATGTGAATCCCGCGGCTTAACTGGGAAAT 600
QY 601 TCCGTTTGAACACAAAGGCTAAGTGTGAGGAGAGAGGAGTGGAAATTCATGTGTAGCAG 660
DB 601 TCCGTTTGAACACAAAGGCTAAGTGTGAGGAGAGAGGAGTGGAAATTCATGTGTAGCAG 660
QY 661 TGAATATGCTAAGATATGGAAGAAATCATGATGGGAGGAGGCTCTCCGTGGTTAACT 720
DB 661 TGAATATGCTAAGATATGGAAGAAATCATGATGGGAGGAGGCTCTCCGTGGTTAACT 720
QY 721 GACGCTCATGACGAAGAGCGTGGGGAGCAAAACAGATTAGATACCTGTTAGTCCACGCG 780
DB 721 GACGCTCATGACGAAGAGCGTGGGGAGCAAAACAGATTAGATACCTGTTAGTCCACGCG 780
QY 781 CTAAACGATGTCAACTAGTGTGGGCTTATTAAGGCTTGTATCGAACGATCAACGCTGA 840
DB 781 CTAAACGATGTCAACTAGTGTGGGCTTATTAAGGCTTGTATCGAACGATCAACGCTGA 840
QY 841 AGTTGACCGCTGGGGAGTACGCTGCGCAAGATTTAAAACTCAAGGAATTTGACGGGGACCC 900
DB 841 AGTTGACCGCTGGGGAGTACGCTGCGCAAGATTTAAAACTCAAGGAATTTGACGGGGACCC 900
QY 901 GACAAAGCGGTGATTAATGATGATTAAATCGATGCAAGCGGAAAACTTACCTACCTT 960
DB 901 GACAAAGCGGTGATTAATGATGATTAAATCGATGCAAGCGGAAAACTTACCTACCTT 960
QY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTCCGGAACGCTTAACACAGGTG 1017
DB 961 GACATGTAGCGAAGCCCGCGAGAGGTGAGTGTGCCGGAAGGAAACGTTAACACAGGTG 1020
QY 1018 CTGCATATGCTGTCTGACGCTGCTGTCTGATGATTTGGGTTAATGTCGCCGCAACGAGCGCA 1077
DB 1021 CTGCATATGCTGTCTGACGCTGCTGTCTGATGATTTGGGTTAATGTCGCCGCAACGAGCGCA 1080
QY 1078 ACCCTTGTGATTAAATTTGCATATTGTTGGGCACTTTAATGAGACCTGCGGTGACAAA 1137
```

```
DB 1081 ACCCTTGTGATTAAATTTGCATATTGCTGGGCACTTTAATGAACTGCCGTGACAAA 1140
QY 1138 CCGAGAGAGGTGGGGATGACGTCAATGCTCTCATAGCCCTTATGGTATGGGCTTACACG 1197
DB 1141 CCGAGAGAGGTGGGGATGACGTCAATGCTCTCATAGCCCTTATGGTATGGGCTTACACG 1200
QY 1198 TAAATCAATGCGGGGTACAGAGGGGTGCGCAACCCCGAGGGGAGACTAATCTCAGAAAGC 1257
DB 1201 TAAATCAATGCGGGGTACAGAGGGGTGCGCAACCCCGAGGGGAGACTAATCTCAGAAAGC 1260
QY 1258 GCGTGTAGTCCGATCCGATCTGCAACTGCACTCCGTGAAGTCGGAATGCTAAT 1317
DB 1261 GCGTGTAGTCCGATCCGATCTGCAACTGCACTCCGTGAAGTCGGAATGCTAAT 1320
QY 1318 CCGGATCAGCATGTCCGCGTGAATACGTTCCCGGCTTGTACACACCCGCCCTGACAC 1377
DB 1321 CCGGATCAGCATGTCCGCGTGAATACGTTCCCGGCTTGTACACACCCGCCCTGACAC 1380
QY 1378 CATGGAGTGGGTTTACACAGACAGGTAGTCTAACCGTAAGAGAGGCGCTTGCCACG 1437
DB 1381 CATGGAGTGGGTTTACACAGACAGGTAGTCTAACCGTAAGAGAGGCGCTTGCCACG 1440
QY 1438 TGAGATTCAATGACTGGGCTG 1457
DB 1441 TGAGATTCAATGACTGGGCTG 1460
```

```
RESULT 38
AY123801
LOCUS
DEFINITION Nitrosospira sp. Nsp12 16S ribosomal RNA gene, partial sequence.
ACCESSION AY123801
VERSION AY123801.1 GI:24474430
KEYWORDS
SOURCE
ORGANISM Nitrosospira sp. Nsp12
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosospira.
REFERENCE
1 (bases 1 to 1497)
Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roeser, A. and
Koops, H.P.
16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial
ammonia-oxidizing isolates: extension of the dataset and proposal
of a new lineage within the nitrosomonads
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)
JOURNAL
PUBMED 13130037
2 (bases 1 to 1497)
Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roeser, A. and
Wagner, M.
Direct Submission
Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES
source
1..1497
/organism="Nitrosospira sp. Nsp12"
/mol_type="genomic DNA"
/strain="Nsp12"
/db_xref="taxon:136545"
<1..>1497
/product="16S ribosomal RNA"
ORIGIN
Query Match 89.6%; Score 1305.2; DB 1; Length 1497;
Best Local Similarity 94.3%; Pred. No. 1.9e-13;
Matches 1377; Conservative 0; Mismatches 78; Indels 5; Gaps 2;
QY 1 ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAACGGCAGACCGATGCTTGCAT 60
DB 3 ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAACGGCAGACCGGCGCAAC-- 60
QY 61 CTGTGGCGAGTGGCGGACGGGTGAGTAAATGCAATCGAACCTATCCAGAAAGAGGGGGTGA 120
```

```

Db      61 CTGTGGCGAGTGGCGAACGGGTGAGTAAATGCAATCGGAACGTAATCTTAAATGGGGGATA 120
Qy      121 AGCATCGAAGAGTGTCTAATATACCGCATATATCTTAAGAGAGAAAGCAGGGAGATCGAAA 180
Db      121 AGCGATCGAAGAGTGTCTAATATACCGCATATATCTTAAGAGAGAAAGCAGGGAGATCGAAA 180
Qy      181 GACCTTGGCGCTTTTGGAGCGGCCGATGTCTGATTAGCTAGTGGTGGGTAAGGCTTAC 240
Db      181 GACCTTGGCGCTTTTGGAGCGGCCGATGTCTGATTAGCTAGTGGTGGGTAAGGCTTAC 240
Qy      241 CAAGGGAAGATCAAGTATGTGCTGAGAGAGACGACGACCACTGGGAGCTGAGACAG 300
Db      241 CAAGGGAAGATCAAGTATGTGCTGAGAGAGACGACGACCACTGGGAGCTGAGACAG 300
Qy      301 GCCCAGACTCTTACGCGGAGGACAGCAGTGGGGAAATTTTGGACAATGGGCGCAACCTGATC 360
Db      301 GCCCAGACTCTTACGCGGAGGACAGCAGTGGGGAAATTTTGGACAATGGGCGCAACCTGATC 360
Qy      361 CAGCAATGCCGCGTGAAGTGAAGAGGCCCTTCCGGTGTAAAGCTCTTTCAGTCAAGAGA 420
Db      361 CAGCAATGCCGCGTGAAGTGAAGAGGCCCTTCCGGTGTAAAGCTCTTTCAGTCAAGAGA 420
Qy      421 AAAAGTTACGTTAAATATATCTGATCTCATGACGGTATCGACAGAAAGAACCGGCTAAC 480
Db      421 AAAAGTTACGTTAAATATATCTGATCTCATGACGGTATCGACAGAAAGAACCGGCTAAC 480
Qy      481 TAGCTGCCAGACGCCGCGGTATACGAGGTGCAAGCGTTAATCGGAATTAATCTGGGCGT 540
Db      481 TAGCTGCCAGACGCCGCGGTATACGAGGTGCAAGCGTTAATCGGAATTAATCTGGGCGT 540
Qy      541 AAAAGGTGCGCAGCGCGCTTTGTAAATCAGATGTGAATCCCGGGCTTAACTTGAGAT 600
Db      541 AAAAGGTGCGCAGCGCGCTTTGTAAATCAGATGTGAATCCCGGGCTTAACTTGAGAT 600
Qy      601 TGGCTTTGAAATCTAACAAGGTGTAAGTGTGCGACAGAGGAGTGGAAATTCATGTGTACG 660
Db      601 TGGCTTTGAAATCTAACAAGGTGTAAGTGTGCGACAGAGGAGTGGAAATTCATGTGTACG 660
Qy      661 TGAATATGCGTAGATATGGAAGAACATCGATGGCGGACAGCCTCTGGGTTAACTACT 720
Db      661 TGAATATGCGTAGATATGGAAGAACATCGATGGCGGACAGCCTCTGGGTTAACTACT 720
Qy      721 GACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAATACCTGTGATGTCACGCC 780
Db      721 GACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAATACCTGTGATGTCACGCC 780
Qy      781 CTAAACGATGTCAATAGTGTGGGCTTATTAAGCTTGTGTAACGAAGCTTAACGCGTGA 840
Db      781 CTAAACGATGTCAATAGTGTGGGCTTATTAAGCTTGTGTAACGAAGCTTAAACGCGTGA 840
Qy      841 AGTTGACCGCTGGGGAGTAGCGGTGCGAAGATTAAACTCAAGAGATTGACGGGACCC 900
Db      841 AGTTGACCGCTGGGGAGTAGCGGTGCGAAGATTAAACTCAAGAGATTGACGGGACCC 900
Qy      901 GCAACAAGCGGTGATTAATGTGATTAATTCGATGCAACGGGAAACCTTAACCTACCTT 960
Db      901 GCAACAAGCGGTGATTAATGTGATTAATTCGATGCAACGGGAAACCTTAACCTACCTT 960
Qy      961 GACATGTAGGAAATTTTCTAGAGATTAATGTC---TTCCGGAAGCTTAACACAGGTG 1017
Db      961 GACATGTAGGAAATTTTCTAGAGATTAATGTC---TTCCGGAAGCTTAACACAGGTG 1017
Qy      1018 CTGATATGCGTGTGCTGAGCTGCTGCTGATGAGATTTGAGTTAAAGTCCCGACAGAGCGCA 1077
Db      1021 CTGATATGCGTGTGCTGAGCTGCTGCTGATGAGATTTGAGTTAAAGTCCCGACAGAGCGCA 1080
Qy      1078 ACCCTTGTCAATTAATTCGATCATTTGTTGGGCACTTTAATGAGACTGCCGCTGACAAA 1137
Db      1081 ACCCTTGTCAATTAATTCGATCATTTGTTGGGCACTTTAATGAAACTGCCGCTGACAAA 1140
Qy      1138 CCGGAGGAAGGTGGGGATGAGCTCAAGTCTCATAGGCCCTTATAGGGTGAAGGCTTACACG 1197
Db      1141 CCGGAGGAAGGTGGGGATGAGCTCAAGTCTCATAGGCCCTTATAGGGTGAAGGCTTACACG 1200

```

```

Qy      1198 TAATCAATGCGCGCTACAGAGCGTTGGCAACCGCGAGGGGGAGCTAATCTCAGAAAGC 1257
Db      1201 TAATCAATGCGCGCTACAGAGCGTTGGCAACCGCGAGGGGGAGCTAATCTCAGAAAGC 1260
Qy      1258 GCGTGTAGTCCGATCGGAAGTCTGCAACTGACTCGCTGAAATGCGAATGCTAGTAAT 1317
Db      1261 GCGTGTAGTCCGATCGGAAGTCTGCAACTGACTCGCTGAAATGCGAATGCTAGTAAT 1320
Qy      1318 CCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGGCTTGTATCAACACCGCCGTCACAC 1377
Db      1321 CCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGGCTTGTATCAACACCGCCGTCACAC 1380
Qy      1378 CATGGAGTGGGTTTACCAAGAGAGATGATCTTAACGTTAAGAGAGGGGCGCTTGCACAG 1437
Db      1381 CATGGAGTGGGTTTACCAAGAGAGATGATCTTAACGTTAAGAGAGGGGCGCTTGCACAG 1440
Qy      1438 TGAGATTCAATGACTGGGGTG 1457
Db      1441 TGAGATTCAATGACTGGGGTG 1460

```

```

RESULT 39
AY123809      1497 bp      DNA      linear      BCT 23-SEP-2003
LOCUS        Nitrosospira sp. II17 16S ribosomal RNA gene, partial sequence.
DEFINITION   AY123809
ACCESSION    AY123809
VERSION      AY123809.1 GI:24474438
KEYWORDS
SOURCE
ORGANISM     Nitrosospira sp. II17
              Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
              Nitrosomonadaceae; Nitrosospira.

```

```

REFERENCE    1 (bases 1 to 1497)
              Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roeser,A. and
              Koops,H.P.
              16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial
              ammonia-oxidizing isolates: extension of the dataset and proposal
              of a new lineage within the nitrosomonade
              Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)

```

```

JOURNAL      13130037
PUBMED
REFERENCE    2 (bases 1 to 1497)
              Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Roeser,A. and
              Wagner,M.
              Direct Submission
              Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische
              Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany

```

```

FEATURES
  source
    1..1497
      /organism="Nitrosospira sp. II17"
      /mol_type="genomic DNA"
      /strain="II17"
      /db_xref="taxon:174932"
      <1..>1497
      /product="16S ribosomal RNA"

```

## ORIGIN

Query Match 89.5%; Score 1303.6; DB 1; Length 1497;

Best Local Similarity 94.2%; Pred. No. 2e-13;

Matches 1376; Conservative 0; Mismatches 79; Indels 5; Gaps 2;

```

Qy      1 ATTGAACGCTGGCGGATCGCTTTTACATGCAAGTGAACCGGACAGCGGATCTTGAT 60
Db      3 ATTGAACGCTGGCGGATCGCTTTTACATGCAAGTGAACCGGACAGCGGATCTTGAT 60
Qy      61 CTGTGGCGAGTGGCGGAGCGGGTGAATGATCGGAACGTAATCCAGAAAGAGGGGGGTA 120
Db      61 CTGTGGCGAGTGGCGGAGCGGGTGAATGATCGGAACGTAATCCAGAAAGAGGGGGGTA 120
Qy      121 ACCGATCGAAGAGTGTCTAATATCCGATATATCTTAAGAGAGAAAGCAGGGGATGAAA 180
Db      121 ACCGATCGAAGAGTGTGGCTTAATACCGCATATATCTTGAAGAGAAAGCAGGGGATGAAA 180

```



```

QY 181 GACCTTGGCTTTGGAGGCGCGATGCTGATTAAGCTAGTTGTTGGGCTAAAGGCTTAC 240
    |||
Db 181 GACCTTGGCTTTGGAGGCGCGATGCTGATTAAGCTAGTTGTTGGGCTAAAGGCTTAC 240
QY 241 CAAGGCGACGATCAGTAGTGGTCTGAGAGAGACACACGACCACTGGGACTGAGACAG 300
    |||
Db 241 CAAGGCTTGCATCAGTAGTGGTCTGAGAGAGACACACGACCACTGGGACTGAGACAG 300
QY 301 GCCCAGACTCTTAACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
    |||
Db 301 GCCCAGACTCTTAACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 361 CAGCATGCGCGGTGAGTGAAGAGGCGCTTGGGCTTGAAGCTCTTTCACTGAGAGAGA 420
    |||
Db 361 CAGCATGCGCGGTGAGTGAAGAGGCGCTTGGGCTTGAAGCTCTTTCACTGAGAGAGA 420
QY 421 AAAGGTTACGCTAAATTAATCTGACTCATGACGCTATCGACAGAAAGACACGCGCTAC 480
    |||
Db 421 AAAGGTTACGCTAAATTAATCTGACTCATGACGCTATCGACAGAAAGACACGCGCTAC 480
QY 481 TACGTTGCCAGACCGCGGCTTAATACGTTAGGCTGCAAGCTTAACTGGAGCT 540
    |||
Db 481 TACGTTGCCAGACCGCGGCTTAATACGTTAGGCTGCAAGCTTAACTGGAGCT 540
QY 541 AAAGGTTGCGAGCGGCTTTGTAAGTCAAGTGTGAATCCCGGCTTAACTGGAGCT 600
    |||
Db 541 AAAGGTTGCGAGCGGCTTTGTAAGTCAAGTGTGAATCCCGGCTTAACTGGAGCT 600
QY 601 TCCGTTTGAACACTCAAGGCTTGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
    |||
Db 601 TCCGTTTGAACACTCAAGGCTTGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 TGAATTCGCTAAGATATGAAAGAACTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
    |||
Db 661 TGAATTCGCTAAGATATGAAAGAACTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 721 GACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
    |||
Db 721 GACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
QY 781 CTAAACGATGTCATAGTGTGTTGGGCTTATTAAGGCTTGTGTAACGACGCTAACCGGCTA 840
    |||
Db 781 CTAAACGATGTCATAGTGTGTTGGGCTTATTAAGGCTTGTGTAACGACGCTAACCGGCTA 840
QY 841 AGTTGACCGCTGGGAGGAGTACGCTGCGAAGATTAATACTCAAGAACTGACGAGGAGCC 900
    |||
Db 841 AGTTGACCGCTGGGAGGAGTACGCTGCGAAGATTAATACTCAAGAACTGACGAGGAGCC 900
QY 901 GCAACAAGCGGTGATTAATGATGATTAATGATGATTAATGATGATTAATGATGATTAAT 960
    |||
Db 901 GCAACAAGCGGTGATTAATGATGATTAATGATGATTAATGATGATTAATGATGATTAAT 960
QY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTGCGGAAACGCTTAAACACAGGTG 1017
    |||
Db 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTGCGGAAACGCTTAAACACAGGTG 1017
QY 1018 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077
    |||
Db 1018 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077
QY 1021 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
    |||
Db 1021 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1078 ACCCTTGTCAATTAATGATGATTAATGATGATTAATGATGATTAATGATGATTAATGAT 1137
    |||
Db 1078 ACCCTTGTCAATTAATGATGATTAATGATGATTAATGATGATTAATGATGATTAATGAT 1137
QY 1081 ACCCTTGTCAATTAATGATGATTAATGATGATTAATGATGATTAATGATGATTAATGAT 1140
    |||
Db 1081 ACCCTTGTCAATTAATGATGATTAATGATGATTAATGATGATTAATGATGATTAATGAT 1140
QY 1138 CCGGAGGAAAGTGGGAGTACGCTCAAGTCTCAATGCGCTTATGGGTAGGGCTTCAACAG 1197
    |||
Db 1138 CCGGAGGAAAGTGGGAGTACGCTCAAGTCTCAATGCGCTTATGGGTAGGGCTTCAACAG 1197
QY 1141 CCGGAGGAAAGTGGGAGTACGCTCAAGTCTCAATGCGCTTATGGGTAGGGCTTCAACAG 1200
    |||
Db 1141 CCGGAGGAAAGTGGGAGTACGCTCAAGTCTCAATGCGCTTATGGGTAGGGCTTCAACAG 1200
QY 1198 TAAATCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
    |||
Db 1198 TAAATCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
QY 1201 TAAATCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
    |||
Db 1201 TAAATCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1258 GGGTGTGATCCGGATCGGAGTCTGCAACTCGACTCGCTGGAAGTGGGAATCGCTAGTAAT 1317
    |||

```

```

Db 1261 GGGTGTGATCCGGATCGGAGTCTGCAACTGCACTCGCTGAAGTCGGAATCGTAGTAAT 1320
    |||
QY 1318 CGCGGATCAGCATGTCGCGGTGATTAATGCTTCCGGGCTTTTATACACACCGCCCTGACAC 1377
    |||
Db 1321 CGCGGATCAGCATGTCGCGGTGATTAATGCTTCCGGGCTTTTATACACACCGCCCTGACAC 1380
    |||
QY 1378 CATGGAGTGGGTTTACACAGAGAGGATGCTTAACCGTAAAGAGGAGGAGGAGGAGGAGGAG 1437
    |||
Db 1381 CATGGAGTGGGTTTACACAGAGAGGATGCTTAACCGTAAAGAGGAGGAGGAGGAGGAGGAG 1440
    |||
QY 1438 TGAGATTCATGACTGGGGTG 1457
    |||
Db 1441 TGAGATTCATGACTGGGGTG 1460
    |||

RESULT 40
AY631270 1458 bp DNA linear BCT 28-DEC-2004
LOCUS Nitrosovibrio sp. FJ1423 16S ribosomal RNA gene, partial sequence.
DEFINITION AY631270
ACCESSION AY631270.1 GI:51093365
VERSION
KEYWORDS
SOURCE
ORGANISM Nitrosovibrio sp. FJ1423
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosovibrio.
REFERENCE
1 (bases 1 to 1458)
Ida,T., Seton,M., Yabe,R., Takahashi,R. and Tokuyama,T.
Identification of Genus Nitrosovibrio, Ammonia-Oxidizing Bacteria,
by Comparison of N-Terminal Amino Acid Sequences of
Phosphoglycerate Kinase
J. Biosci. Bioeng. 98 (5), 380-383 (2004)
2 (bases 1 to 1458)
Takahashi,R., Yabe,R. and Tokuyama,T.
Nitrosovibrio sp. FJ1423 16S ribosomal RNA gene
Unpublished
3 (bases 1 to 1458)
Takahashi,R., Yabe,R. and Tokuyama,T.
Direct Submission
Submitted (20-MAY-2004) Agricultural and Biological Chemistry,
College of Bioresource Sciences, Nihon University, 1866 Kameino,
Fujisawa, Kanagawa 252-8510, Japan
FEATURES
source location/Qualifiers
1..1458
/organism="Nitrosovibrio sp. FJ1423"
/mol_type="genomic DNA"
/db_xref="taxon:289540"
/clone="FJ1423"
<1..>1458
/product="16S ribosomal RNA"

ORIGIN
Query Match 89.4%; Score 1302; DB 1; Length 1458;
Best Local Similarity 94.2%; Pred. No. 2.2e-13;
Matches 1375; Conservative 0; Mismatches 80; Indels 5; Gaps 2;

QY 1 ATGGAACGCTGGCGGATGCTTACATGCAATGCAAGTCGAAACGACGAGGAGGAGGAGGAGGAG 60
    |||
Db 1 ATGGAACGCTGGCGGATGCTTACATGCAATGCAAGTCGAAACGACGAGGAGGAGGAGGAGGAG 58
    |||
QY 61 CTGGTGGAGTGGGCGGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 120
    |||
Db 59 CTGGTGGAGTGGGCGGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 118
    |||
QY 121 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGAGGATCGAAA 180
    |||
Db 119 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGAGGATCGCAA 178
    |||
QY 181 GACCTTGGCTTTTGGAGGCGCGCATGCTGATTAAGCTAGTTGTTGGGCTAAAGGCTTAC 240
    |||
Db 179 GACCTTGGCTTTTGGAGGCGCGCATGCTGATTAAGCTAGTTGTTGGGCTAAAGGCTTAC 238
    |||

```



QY 241 CAAAGCGAGATCAGTGGTCTGAGAGACGACCGCACTGGGACCTGAGACACG 300  
 DB 229 CAAAGCGAGATCAGTGGTCTGAGAGAGACGACCGCACTGGGACCTGAGACACG 298  
 QY 301 GCCCGAGCTCTTACGCGGAGGCGAGCTGGGGAAATTTTGGACAATGGGCGCAAGCCTGATC 360  
 DB 299 GCCCGAGCTCTTACGCGGAGGCGAGCTGGGGAAATTTTGGACAATGGGCGCAAGCCTGATC 358  
 QY 361 CAGCAATGCCGCGTGAAGTGAAGAGCCCTTCGGGTTTAAAGCTCTTTTCAGTCGAGAGA 420  
 DB 359 CAGCAATGCCGCGTGAAGTGAAGAGCCCTTCGGGTTTAAAGCTCTTTTCAGTCGAGAGA 418  
 QY 421 AAAAGTTACGTTAATATCGTGAATCATGACGGTATGCAAGAGAGACACCGGTTAC 480  
 DB 419 AACGGTTACGTTAATATCGTGAATCATGACGGTATGCAAGAGAGACACCGGTTAC 478  
 QY 481 TACGTTACGAGACCGCGGTTAATACCTAGGGTGCAGCGTTAATCGGAATTAATCTGGGCGCT 540  
 DB 479 TACGTTACGAGACCGCGGTTAATACCTAGGGTGCAGCGTTAATCGGAATTAATCTGGGCGCT 538  
 QY 541 AAAAGGTTGCGAGCGCGCTTTTGAAGTCAAGATGTAATCCCGGGCTTAACTGGGAAT 600  
 DB 539 AAAAGGTTGCGAGCGCGCTTTTGAAGTCAAGATGTAATCCCGGGCTTAACTGGGAAT 598  
 QY 601 TGGGTTGAATCAAGGCTAGAGTGTGGCAGAGGAGTGAATTCATGTGTACAG 660  
 DB 599 TGGGTTGAATCAAGGCTAGAGTGTGGCAGAGGAGTGAATTCATGTGTACAG 658  
 QY 661 TGAATATGCTGAAGATATGGAAGAACATGATGCGAAGGACCGCTCTGGGTTAATCACT 720  
 DB 659 TGAATATGCTGAAGATATGGAAGAACATGATGCGAAGGACCGCTCTGGGTTAATCACT 718  
 QY 721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATAGTACCTGTGATGTCACGCGC 780  
 DB 719 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATAGTACCTGTGATGTCACGCGC 778  
 QY 781 CTAAACGATGTCAACTAGTGTGGGCTTATTAGCTGGTGAACGAGCTTAACGCGTGA 840  
 DB 779 CTAAACGATGTCAACTAGTGTGGGCTTATTAGCTGGTGAACGAGCTTAACGCGTGA 838  
 QY 841 AGTTGACCGCTGGGAGTACCGTCCGAGATTTAAACTAAAGAAATTTGACGGGAGACC 900  
 DB 839 AGTTGACCGCTGGGAGTACCGTCCGAGATTTAAACTAAAGAAATTTGACGGGAGACC 898  
 QY 901 GCAACAGCGGTGAATATGATGATTAATGATGCAACGCGAAACCTTACCTACCTT 960  
 DB 899 GCAACAGCGGTGAATATGATGATTAATGATGCAACGCGAAACCTTACCTACCTT 958  
 QY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC--TTCCGGGAACGCTTAACAGAGTG 1017  
 DB 959 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC--TTCCGGGAACGCTTAACAGAGTG 1018  
 QY 1018 CTGATGAGCTGTGCTGACGCTGCTGTGCTGAGATTTTGGGTTAATCCCCGACGAGGCGCA 1077  
 DB 1019 CTGATGAGCTGTGCTGACGCTGCTGTGCTGAGATTTTGGGTTAATCCCCGACGAGGCGCA 1078  
 QY 1078 ACCCTGTCAATTAATTTGCAATCATTTGGTGGGCACTTTATGAGACTGCGCGGTGACAA 1137  
 DB 1079 ACCCTGTCAATTAATTTGCAATCATTTGGTGGGCACTTTATGAGACTGCGCGGTGACAA 1138  
 QY 1138 CCGGAGAGAGGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACG 1197  
 DB 1139 CCGGAGAGAGGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACG 1198  
 QY 1198 TAAATCAATGAGCGGCTGACAGAGGTTGCAACCCCGAGAGGAGGCTTAATTCGAAAGC 1257  
 DB 1199 TAAATCAATGAGCGGCTGACAGAGGTTGCAACCCCGAGAGGAGGCTTAATTCGAAAGC 1258  
 QY 1258 GCGGTGATGTCGGAATGGAATCTGCACTCCGTTGAAGTCCGAATGCTAGTAAAT 1317  
 DB 1259 GCGGTGATGTCGGAATGGAATCTGCACTCCGTTGAAGTCCGAATGCTAGTAAAT 1318  
 QY 1318 CGCGATACGATGTCGCGGTGAATAGCTTCCCGGATCTTGTACACACCGCCGCTCACAC 1377

DB 1319 CGCGATACGATGTCGCGGTGAATAGCTTCCGGGTCTTGTACACACCGCCCGTCACAC 1378  
 QY 1378 CATGGAGTGGGTTTACCAAGAGAGTACTCTAACCTTAAGAGAGGCGCTTGGCAGCG 1437  
 DB 1379 CATGGAGTGGGTTTACCAAGAGAGTACTCTAACCTTAAGAGAGGCGCTTGGCAGCG 1438  
 QY 1438 TGAGATTACATGCTGGGGTG 1457  
 DB 1439 TGAGATTACATGCTGGGGTG 1458  
 RESULT 41  
 AY631269  
 LOCUS AY631269 1458 bp DNA linear BCT 28-DEC-2004  
 DEFINITION Nitrosovibrio sp. FJ182 16S ribosomal RNA gene, partial sequence.  
 ACCESSION AY631269  
 VERSION AY631269.1 GI:51093364  
 KEYWORDS  
 SOURCE Nitrosovibrio sp. FJ182  
 ORGANISM Nitrosovibrio sp. FJ182  
 Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.  
 REFERENCE  
 AUTHORS Ida, T., Satoh, M., Yabe, R., Takahashi, R. and Tokuyama, T.  
 TITLE Identification of N-Terminal Amino Acid Sequences of Phosphoglycerate Kinase  
 JOURNAL J. Biosci. Bioeng. 98 (5), 380-383 (2004)  
 REFERENCE  
 AUTHORS J. Biosci. Bioeng. 98 (5), 380-383 (2004)  
 TITLE Nitrosovibrio sp. FJ182 16S ribosomal RNA gene  
 JOURNAL Unpublished  
 REFERENCE  
 AUTHORS 3 (bases 1 to 1458)  
 TITLE Takahashi, R., Sato, N. and Tokuyama, T.  
 JOURNAL Direct Submission  
 TITLE Submitted (20-MAY-2004) Agricultural and Biological Chemistry, College of Bioresource Sciences, Nihon University, 1866 Kameino, Fujisawa, Kanagawa 252-8510, Japan  
 FEATURES  
 source  
 1..1458  
 /organism="Nitrosovibrio sp. FJ182"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:289539"  
 /clone="FJ182"  
 <1..>1458  
 /product="16S ribosomal RNA"  
 ORIGIN  
 Query Match 89.0%; Score 1297.2; DB 1; Length 1458;  
 Best local Similarity 94.0%; Pred. No. 2, 6e-13;  
 Matches 1372; Conservative 0; Mismatches 83; Indels 5; Gaps 2;  
 QY 1 ATTGAACGCTGGCGGATGCTTTTACATGACATGCAAGTGAACGCGAGCAGGATCTTGCAT 60  
 DB 1 ATTGAACGCTGGCGGATGCTTTTACATGACATGCAAGTGAACGCGAGCAGGATCTTGCAT 58  
 QY 61 CTGGTGGCAGTGGCGGACGCGGTGAGTAAATGATCGGAAAGTATCCAAAGAGGGGGGTA 120  
 DB 59 CTGGTGGCAGTGGCGGACGCGGTGAGTAAATGATCGGAAAGTATCCAAAGAGGGGGGTA 118  
 QY 121 ACCGATCGAAAGATGTGCTAATACCGCTATATCTTAAGAGAGAAAGCGGGGATGAAA 180  
 DB 119 ACCGATCGAAAGATGTGCTAATACCGCTATATCTTAAGAGAGAAAGCGGGGATGAAA 178  
 QY 181 GACCTTGGCGCTTTTGGAGCGGCGGATGCTGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 240  
 DB 179 GACCTTGGCGCTTTTGGAGCGGCGGATGCTGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 238  
 QY 241 CAAAGCGACGATGATGTTGTCTGAGAGAGACGACGACCACTGGGACTGAGACACG 300  
 DB 239 CAAAGCGACGATGATGTTGTCTGAGAGAGACGACGACCACTGGGACTGAGACACG 298

QY 301 GCCCAGACTCCTACGGGAGGCGAGAGTGGGGAAATTTTGGACAAATGGGGCGACCGCGATC 360  
DB 299 GCCCAGACTCCTACGGGAGGCGAGAGTGGGGAAATTTTGGACAAATGGGGAAACCTTGATC 358  
QY 361 CAGCAATGCCCGGTAGTGAAGAAAGCCCTTCGGGTGTAAAGCTCTTTCACTCGAAGAA 420  
DB 359 CAGCAATGCCCGGTAGTGAAGAAAGCCCTTCGGGTGTAAAGCTCTTTCACTCGAAGAA 418  
QY 421 AAAAGTTACGGTAAATATCTGACTCATGACGGTATCGACAGAAAGAACCCGGCTAAC 480  
DB 419 AACGGTTACGGTAAATATCTGACTCATGACGGTATCGACAGAAAGAACCCGGCTAAC 478  
QY 481 TAACGGTACGAGCCCGGTATAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
DB 479 TAACGGTACGAGCCCGGTATAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 538  
QY 541 AAAAGGTTACGAGCGGCTTTGTAAAGTCAAGTGTAAATCCCGGGCTTAACTTGGGAAT 600  
DB 539 AAAAGGTTACGAGCGGCTTTGTAAAGTCAAGTGTAAATCCCGGGCTTAACTTGGGAAT 598  
QY 601 TCCGTTTGAACCTACAGAGGCTAGAGTGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 599 TCCGTTTGAACCTACAGAGGCTAGAGTGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 658  
QY 661 TGAATTCGCTAGAGTATGGAAGAAACATGATGCGGAAGGCGCTCTCGGTGTTAACT 720  
DB 659 TGAATTCGCTAGAGTATGGAAGAAACATGATGCGGAAGGCGCTCTCGGTGTTAACT 718  
QY 721 GACGCTATGACAGAAAGCGTGGGAGCAGAAACAGATTAGATACCTGTTAGTCCAGCGC 780  
DB 719 GACGCTATGACAGAAAGCGTGGGAGCAGAAACAGATTAGATACCTGTTAGTCCAGCGC 778  
QY 781 CTTAAACGATGTCACTAGTGTGGGCTTATTAAGCTTTGTATGACGAACCTAACCGGTGA 840  
DB 779 CTTAAACGATGTCACTAGTGTGGGCTTATTAAGCTTTGTATGACGAACCTAACCGGTGA 838  
QY 841 AGTTGACCGCTTCGGGAGTACCGGTGCGCAAGATTAAACCTCAAGAAATTGACGGGAGACC 900  
DB 839 AGTTGACCGCTTCGGGAGTACCGGTGCGCAAGATTAAACCTCAAGAAATTGACGGGAGACC 898  
QY 901 GCACAGACGGGTGATATGATGATTAATTCATGCAACCGCAAAACCTTACCTACCTT 960  
DB 899 GCACAGACGGGTGATATGATGATTAATTCATGCAACCGCAAAACCTTACCTACCTT 958  
QY 961 GACATGTACGGAATTTTCTAGAGATAGATTAGTGC--TTCCGGAAACGCTAACACAGGTG 1017  
DB 959 GACATGTACGGAATTTTCTAGAGATAGATTAGTGC--TTCCGGAAACGCTAACACAGGTG 1018  
QY 1018 CTGCAATGCTGTCTGCTCACTGCTGTCTGTGATGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1077  
DB 1019 CTGCAATGCTGTCTGCTCACTGCTGTCTGTGATGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1078  
QY 1078 ACCCTTGTCAATTAATTTGACATTTGTTGGGCACTTAAATGAGACTCCGGTACAAA 1137  
DB 1079 ACCCTTGTCAATTAATTTGACATTTGTTGGGCACTTAAATGAGACTCCGGTACAAA 1138  
QY 1138 CCGGAGGAAGTGGGGATGACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACAG 1197  
DB 1139 CCGGAGGAAGTGGGGATGACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACAG 1198  
QY 1198 TAAATCAATGCGCGCTACAGAGGTTGCGCAACCCCGAGGGGAGGATTAATCTCAAGAAAC 1257  
DB 1199 TAAATCAATGCGCGCTACAGAGGTTGCGCAACCCCGAGGGGAGGATTAATCTCAAGAAAC 1258  
QY 1258 GCGTGTAGTCCGAGTCCGAGTCTGCACTCCGTAACCTCGGAATCGCTAGTAAT 1317  
DB 1259 GCGTGTAGTCCGAGTCCGAGTCTGCACTCCGTAACCTCGGAATCGCTAGTAAT 1318  
QY 1318 CCGGATCAAGATGTCGCGGTAAATACGTTCCGGGTCTTGTACACACCGCCGCTACAC 1377  
DB 1319 CCGGATCAAGATGTCGCGGTAAATACGTTCCGGGTCTTGTACACACCGCCGCTACAC 1378  
QY 1378 CATGGAGTGGGTTTCAACCAAGACAGTATGCTAACCGTAAAGAGGGCGCTTGGCACAGG 1437

DB 1379 CATGGAGTGGGTTTCAACCAAGACAGTATCTTAACCGCAAGAGGGCGCTTGGCACAGG 1438  
QY 1438 TGAATTCATGACTGGGGTGG 1457  
DB 1439 TGAATTCATGACTGGGGTGG 1458  
RESULT 42  
A1621028 1462 bp DNA linear BCT 08-JUN-2004  
LOCUS Nitrosomonas sp. 1686 16S rRNA gene, isolate 1686.  
DEFINITION A1621028  
ACCESSION A1621028  
VERSION A1621028.1 GI:40994847  
KEYWORDS 16S ribosomal RNA; 16S rRNA gene.  
SOURCE Nitrosomonas sp. 1686  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1  
AUTHORS Nielsen, S., Baer Gillissen, M. U., Revebeck, N., Laanbroek, H. J. and Bollmann, A.  
TITLE Influence of heterotrophic bacteria on the activity and growth behavior of ammonia-oxidizing bacteria  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1462)  
AUTHORS Bollmann, A.  
TITLE Direct Submersion  
JOURNAL Submitted (15-JAN-2004) Bollmann A., Department for Microbial Ecology, NIOO-KNAW Centre for Limnology, Rijksstraatweg 6, 3631 AC Nieuwerstadij, NETHERLANDS  
FEATURES  
source location/Qualifiers  
1..1462  
/organism="Nitrosomonas sp. 1686"  
/mol\_type="genomic DNA"  
/isolate="1686"  
/isolation\_source="fresh water sediment"  
/db\_xref="taxon:261294"  
/country="Netherlands; Lake Drontemeer"  
1..1462  
/gene="16S rRNA"  
1..1462  
/gene="16S rRNA"  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 89.0%; Score 1296.8; DB 1; Length 1462;  
Best Local Similarity 94.6%; Pred. No. 2.6e-13;  
Matches 1375; Conservative 0; Mismatches 73; Indels 5; Gaps 3;  
QY 1 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCCGAACGGCAGCAGATGCTTGCAT 60  
DB 12 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCCGAACGGCAGCAGATGCTTGCAT 71  
QY 61 CTGGTGGGAGTGGCGGAGCGGGTGAATGATCGGAACGATCCGAAGAGGGGGGTA 120  
DB 72 CTGGTGGGAGTGGCGGAGCGGGTGAATGATCGGAACGATCTTAAATGGGGGTA 131  
QY 121 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGAAAGCAGGGGATCGAA 180  
DB 132 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGAAAGTGGGGATCCAAA 191  
QY 181 GACCTTGGCTTTTGGAGACGGCCGATGTCTGATTAAGTGAATTTGGTGGGCTTAAAGGCTTAC 240  
DB 192 AGCCTCATGCTTTTGGAGACGGCCGATGTCTGATTAAGTGAATTTGGTGGGCTTAAAGGCTTAC 251  
QY 241 CAAGCGCAAGTATCGATGTTGTCTGAGAGGAGCAGACCACTGGGACTGAGACAG 300  
DB 252 CAAGCGCAAGTATCGATGTTGTCTGAGAGGAGCAGACCACTGGGACTGAGACAG 311  
QY 301 GCCCAGACTCCTACGGGAGGCGAGAGTGGGGAATTTTGGACAAATGGGCGCAAGCCTGATC 360  
DB 312 GCCCAGACTCCTACGGGAGGCGAGAGTGGGGAATTTTGGACAAATGGGCGCAAGCCTGATC 371

QY 361 CAGCAATGCGCGTAGTGAAGAGCGCTTCGGGTGTAAAGCTCTTTTCAGTCGAGAGA 420  
 DB 372 CAGCAATGCGCGTAGTGAAGAGCGCTTCGGGTGTAAAGCTCTTTTCAGTCGAGAGA 430  
 QY 421 AAAGGTTACGGTAATATCTGATCTCATGACGGTATCGACAGAGAGACACCGGCTAAC 480  
 DB 431 AAAAATTTCTGGCTAATACCAAGATTCATGACGGTATCGACAGAGAGACACCGGCTAAC 490  
 QY 481 TACGTCACGACGCGCGGTAAATACGTAAGGTCGACGCGTAAATTCGGAATTAATCTGGGCGT 540  
 DB 491 TACGTCACGACGCGCGGTAAATACGTAAGGTCGACGCGTAAATTCGGAATTAATCTGGGCGT 550  
 QY 541 AAAGGTCGACGCGCGCTTTGTAACTGATGATGTAATCCCGGCGTTAACTCTGGGAAT 600  
 DB 551 AAAGGTCGACGCGCGCTTTGTAACTGATGATGTAATCCCGGCGTTAACTCTGGGAAT 610  
 QY 601 TGGGTTGAAATCTACAAAGGCTAGAGTGTGCGACAGGAGGTGGAATTTCCATGTGTACAG 660  
 DB 611 TGGGTTGAAATCTACAAAGGCTAGAGTGTGCGACAGGAGGTGGAATTTCCATGTGTACAG 670  
 QY 661 TGAATAGCGTAAGATATGGAAGAATGATGCGGACGACGCTCTGGGTTAACTACT 720  
 DB 671 TGAATAGCGTAAGATATGGAAGAATGATGCGGACGACGCTCTGGGTTAACTACT 730  
 QY 721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGTACTCAAGCC 780  
 DB 731 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGTACTCAAGCC 790  
 QY 781 CTAAACGATGTCAACTAGTTGTTGGGCTTATTAGGCTTGTAAACGAGCTAAACGCGTGA 840  
 DB 791 CTAAACGATGTCAACTAGTTGTTGGGCTTATTAGGCTTGTAAACGAGCTAAACGCGTGA 850  
 QY 841 AGTTGACCGCGTGGGAGTACGCTCGCAAGATTAAATCTAAAGGAATTGACGGGGAGCC 900  
 DB 851 AGTTGACCGCGTGGGAGTACGCTCGCAAGATTAAATCTAAAGGAATTGACGGGGAGCC 910  
 QY 901 GCAACAAGCGGTGATTTATGATGATTAATTCATGATGCAACGCGAAACCTTAACCTACCTT 960  
 DB 911 GCAACAAGCGGTGATTTATGATGATTAATTCATGATGCAACGCGAAACCTTAACCTACCTT 970  
 QY 961 GACATGTAGCGAAATTTCTAGAGATAGATTAGTCT--TCGGGAACGCTTAACACAGGTG 1017  
 DB 971 GACATGTAGCGAAATTTCTAGAGATAGATTAGTCTCGAAAGAGAACTTAACACAGGTG 1030  
 QY 1018 CTGATGAGCTGTCTGCTGAGTCTGTCTGATGATGTTGGGTTAACTCCCGGACGAGCGCA 1077  
 DB 1031 CTGATGAGCTGTCTGCTGAGTCTGTCTGATGATGTTGGGTTAACTCCCGGACGAGCGCA 1090  
 QY 1078 ACCCTGTCAATTAATGCGCATTTGGTGGGCACTTTATGAGACTGCGCGGTGACAA 1137  
 DB 1091 ACCCTGTCAATTAATGCGCATTTAACTTGGGCACTTTATGAGACTGCGCGGTGATTA 1150  
 QY 1138 CCGAGAGAGGTGGGATGACGTCAGTCTCATGAGCCCTTATGGGATGGGCTTACACG 1197  
 DB 1151 CCGAGAGAGGTGGGATGACGTCAGTCTCATGAGCCCTTATGGGATGGGCTTACACG 1210  
 QY 1198 TAAATCAATGCGCGTACAGAGGTTGCGCAACCGCGAGGGGAGCTTAATTCGAAAGC 1257  
 DB 1211 TAAATCAATGCGCGTACAGAGGTTGCGCAACCGCGAGGGGAGCTTAATTCGAAAGC 1270  
 QY 1258 GCGTCGATGTCGCGATGCGAGTCTGCAACTGCACTCGGTGAAGTCGGAATCGCTAGTAAT 1317  
 DB 1271 GCGTCGATGTCGCGATGCGAGTCTGCAACTGCACTCGGTGAAGTCGGAATCGCTAGTAAT 1330  
 QY 1318 CGCGGATCAGCATGTGCGGTGAATACGTTCCGGGCTTTGTACACACCGCGCTCAAC 1377  
 DB 1331 CGCGGATCAGCATGTGCGGTGAATACGTTCCGGGCTTTGTACACACCGCGCTCAAC 1390  
 QY 1378 CATGGAGTGGGTTTACCAAGAGCATGATCTTAACGTTAAGAGGCGGCTTTCACACG 1437  
 DB 1391 CATGGAGTGGGTTTACCAAGAGCATGATCTTAACGTTAAGAGGCGGCTTTCACACG 1449

QY 1438 TGAGATTCATGAC 1450  
 DB 1450 TGAGATTCATGAC 1462  
 RESULT 43  
 AY856079  
 LOCUS Nitrosospira sp. En13 1495 bp DNA linear BCT 12-JUN-2005  
 DEFINITION Nitrosospira sp. En13 16S ribosomal RNA gene, partial sequence.  
 ACCESSION AY856079  
 VERSION AY856079.1 GI:57232110  
 KEYWORDS  
 SOURCE  
 ORGANISM Nitrosospira sp. En13  
 Nitrosospira sp. En13  
 Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
 Nitrosomonadaceae; Nitrosospira.  
 REFERENCE 1 (bases 1 to 1495)  
 Shaw, L.J., Nicol, G.W., Smith, Z., Fear, J., Prosser, J.I. and  
 Baggs, E.M. Nitrous oxide production and nitrifier denitrification by  
 betaproteobacterial ammonia oxidising bacteria  
 TITLE Unpublished  
 JOURNAL 2 (bases 1 to 1495)  
 REFERENCE Shaw, L.J., Nicol, G.W., Smith, Z., Fear, J., Prosser, J.I. and  
 Baggs, E.M. Direct Submission  
 AUTHORS Submitted (14-DEC-2004) School of Biological Sciences, University  
 of Aberdeen, Crutchebank Building, St Machar Drive, Aberdeen AB24  
 3UU, U.K.  
 FEATURES  
 source  
 1..1495 Location/Qualifiers  
 /organism="Nitrosospira sp. En13"  
 /mol\_type="genomic DNA"  
 /strain="En13"  
 /db\_xref="taxon:308023"  
 <1..>1495  
 /product="16S ribosomal RNA"  
 ORIGIN  
 Query Match 88.9%; Score 1295.6; DB 1; Length 1495;  
 Best Local Similarity 93.9%; Pred. No. 2,7e-13;  
 Matches 1371; Conservative 0; Mismatches 84; Indels 5; Gaps 2;  
 QY 1 ATTGAACGCTGCGCGCATGCTTTTACATGCAAGTGAACGCGACGACCGGATCTTGCAT 60  
 DB 1 ATTGAACGCTGCGCGCATGCTTTTACATGCAAGTGAACGCGACGACCGGAGCAAC-- 58  
 QY 61 CTGCTGCCAGTGGCGGACCGGCTGATGATGCAATCGGAAAGTATCCAGAAAGAGGGGAT 120  
 DB 59 CTGCTGCCAGTGGCGGACCGGCTGATGATGCAATCGGAAAGTATCTTCGATGGGGAT 118  
 QY 121 AGCATCGAAGAGATGTCATATACCGCATATCTTAAGAGGAGAAAGCAGGGATCGAA 180  
 DB 119 AGCATCGAAGAGATGTCATATACCGCATATCTTAAGAGGAGAAAGCAGGGATCGCA 178  
 QY 181 GACCTTGCGCTTTTGAAGCGCGCGATGCTGATTTAGTATGTTGGTGGGCTTAAGGCTTAC 240  
 DB 179 GACCTTGCGCTTTTGAAGCGCGCGATGCTGATTTAGTATGTTGGTGGGCTTAAGGCTTAC 238  
 QY 241 CAAGGCGACATCATGATGTTGTTGAGAGAGAGACGACGACACCTGGAGCTGAGACAGC 300  
 DB 239 CAAGGCGACATCATGATGTTGTTGAGAGAGAGACGACGACACCTGGAGCTGAGACAGC 298  
 QY 301 GCCAGACTCTCTACGCGGAGCAGACAGTGGGAAATTTTGAACAATGGGCGGACGCTGATC 360  
 DB 299 GCCAGACTCTCTACGCGGAGCAGACAGTGGGAAATTTTGAACAATGGGCGGAAACCTGATC 358  
 QY 361 CAGCAATGCGCGTGAAGTGAAGAGCGCTTTCGGGTTGAAGCTTTTCACTGAGAGAGA 420  
 DB 359 CAGCAATGCGCGTGAAGTGAAGAGCGCTTTCGGGTTGAAGCTTTTCACTGAGAGAGA 418  
 QY 421 AAAGTTACGGTAATATCTGATCTCATGACGGTATCGACAGAGAGACACCGGCTAAC 480

```
Db 419 AACGGTCACGGTTAATACCCGTAAGTACGACGGTACCGAAGAGAACACCGGCTAAC 478
Qy 481 TACGTCCACGACGCGCGGTAACTAGAGGTGCAAGCCTTAATCGAATTACTGGCGCT 540
Db 479 TACGTCCACGACGCGCGGTAACTAGAGGTGCGAGCCTTAATCGAATTACTGGCGCT 538
Qy 541 AAAAGGTGGCAGCGCGCTTTGTAGTCAAGTGTAAATCCCCGGGCTTAACCTGGGAAT 600
Db 539 AAAAGGTGGCAGCGCGCTTTGTAGTCAAGTGTAAATCCCCGGGCTTAACCTGGGAAT 598
Qy 601 TCGGTTTGAACCTACAAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTAGACAG 660
Db 599 TCGGTTTGAACCTACAAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTAGACAG 658
Qy 661 TGAATATGCTAGATATATGAAAGAAATGATGGCGAAGGCGACGCTCTGGGTTAACT 720
Db 659 TGAATATGCTAGATATATGAAAGAAATGATGGCGAAGGCGACGCTCTGGGTTAACT 718
Qy 721 GACGCTCATGCAAGAAAGGTGGGAGCAAGAACTTGATATCCCTGGTAGTCCACGCGC 780
Db 719 GACGCTCATGCAAGAAAGGTGGGAGCAAGAACTTGATATCCCTGGTAGTCCACGCGC 778
Qy 781 CTAAACGATGTCAACTAGTTGTGGGCTTAATTAAGGCTTGTAACGAAGCTAACGCGTGA 840
Db 779 CTAAACGATGTCAACTAGTTGTGGGCTTAATTAAGGCTTGTAACGCGTGA 838
Qy 841 AGTTGACCGCTGGGAGTACGGTGCAGAAATTAATACTCAAGAAATGACGGGAGCC 900
Db 839 AGTTGACCGCTGGGAGTACGGTGCAGAAATTAATACTCAAGAAATGACGGGAGCC 898
Qy 901 GGAAGAGCGGTGGAATTAATGGAATTAATTCGATGCAAGCGGAAACCTTACCTACCT 960
Db 899 GGAAGAGCGGTGGAATTAATGGAATTAATTCGATGCAAGCGGAAACCTTACCTACCT 958
Qy 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTCCGGAACGCTTAACACAGGTG 1017
Db 959 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTCCGGAACGCTTAACACAGGTG 1018
Qy 1018 CTGCATAGGCTGTCTGTCAGTCTGTCTGTGAGATGTTGGGTTAAAGTCCGCAACGAGCGCA 1077
Db 1019 CTGCATAGGCTGTCTGTCAGTCTGTCTGTGAGATGTTGGGTTAAAGTCCGCAACGAGCGCA 1078
Qy 1078 ACCCTTGTCATTAATGTCATATTTGGTGGGCACTTAATGAGACGTCGGCTGACAA 1137
Db 1079 ACCCTTGTCATTAATGTCATATTTGGTGGGCACTTAATGAGACGTCGGCTGACAA 1138
Qy 1138 CCGGAGGAAGGTGGGATGACGTCAAGTCTCTATGGCCCTTAATGGGTAGGGCTTCAACG 1197
Db 1139 CCGGAGGAAGGTGGGATGACGTCAAGTCTCTATGGCCCTTAATGGGTAGGGCTTCAACG 1198
Qy 1198 TAAATCAATGCGCGCTACAGAGGTTGCCAACCCCGGAGGGGAGCTTAATTCAGAAAGC 1257
Db 1199 TAAATCAATGCGCGCTACAGAGGTTGCCAACCCCGGAGGGGAGCTTAATTCAGAAAGC 1258
Qy 1258 GGTGTGTAGTCCGGATCCGAGTCTGCAACTGCACTCCGTAAGTGGGAATGCTATGTAAT 1317
Db 1259 GGTGTGTAGTCCGGATCCGAGTCTGCAACTGCACTCCGTAAGTGGGAATGCTATGTAAT 1318
Qy 1318 CCGGATCAGCATGTCGCGGTAAATACGTTCCCGGGTCTTGTACACACCGCCGCTCACAC 1377
Db 1319 CCGGATCAGCATGTCGCGGTAAATACGTTCCCGGGTCTTGTACACACCGCCGCTCACAC 1378
Qy 1378 CATGGAGTGGGTTTTCACAGAAAGCATGATCTTAAACGTAAGAGAGGCGCTTGCCACG 1437
Db 1379 CATGGAGTGGGTTTTCACAGAAAGCATGATCTTAAACGTAAGAGAGGCGCTTGCCACG 1438
Qy 1438 TGAAGATTCATGACTGGGGTG 1457
Db 1439 TGAAGATTCATGACTGGGGTG 1458
```

RESULT 44  
AY123813

```
LOCUS AY123813 1497 bp DNA linear BCT 23-SEP-2003
DEFINITION Nitrosospira sp. Nsp65 16S ribosomal RNA gene, partial sequence.
ACCESSION AY123813
VERSION AY123813.1 GI:24474442
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
FEATURES
source
1..1497
/organism="Nitrosospira sp. Nsp65"
/mol_type="genomic DNA"
/strain="Nsp65"
/db_xref="taxon:200121"
<1..>1497
/product="16S ribosomal RNA"
ORIGIN
Query Match 88.9%; Score 1295.6; DB 1; Length 1497;
Best Local Similarity 93.9%; Pred. No. 2.7e-13;
Matches 1371; Conservative 0; Mismatches 84; Indels 5; Gaps 2;
Qy 1 ATTGAACGCTGGCGGCAATCTTACATGCAAGTCGAAGCGGACGACGATGCTTCAT 60
Db 3 ATTGAACGCTGGCGGCAATCTTACATGCAAGTCGAAGCGGACGACGATGCTTCAT 60
Qy 61 CTGTGTGGAGTGGCGGACGCGGTGAGTATCATCGAAGCTATTCAGAAAGGGGGTGA 120
Db 61 CTGTGTGGAGTGGCGGACGCGGTGAGTATCATCGAAGCTATTCAGAAAGGGGGTGA 120
Qy 121 ACGCATGAAAGATGTCTAATACCGCATATCTTAAGAGAAAGCAGGGGATCGAAA 180
Db 121 ACGCATGAAAGATGTCTAATACCGCATATCTCTGAGAGAAAGCAGGGGATCGCAA 180
Qy 181 GACCTTGGCTTTTGGACGCGGCCGATGTCATTAAGTATGTTGGTGGGTTAAAGCCTAC 240
Db 181 GACCTTGGCTTTTGGACGCGGCCGATGTCATTAAGTATGTTGGTGGGTTAAAGCCTAC 240
Qy 241 CAAGCGCATGATCACTAGTATGTTGTTGAGAGACGACCACTGGGACTGAGACAG 300
Db 241 CAAGCGCATGATCACTAGTATGTTGTTGAGAGACGACCACTGGGACTGAGACAG 300
Qy 301 GCCCAGATCTCTACGGGAGGCGACAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360
Db 301 GCCCAGATCTCTACGGGAGGCGACAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360
Qy 361 CAGCAATGCGCGTGAAGGAAGAGGCTTGGGGTGAAGCTCTTTCAGTCCGAGAGA 420
Db 361 CAGCAATGCGCGTGAAGGAAGAGGCTTGGGGTGAAGCTCTTTCAGTCCGAGAGA 420
Qy 421 AAAAGTTACGGTAAATATATCGTACTCATGACGATATGACAGAAAGCAACCGGCTAAC 480
Db 421 AAAAGTTACGGTAAATATATCGTACTCATGACGATATGACAGAAAGCAACCGGCTAAC 480
Qy 481 TACGTCCACGACGCGCGGTAAATACGTAAGGTTGCAAGCGTTAATTCGGAATTAATCTGGGCGT 540
Db 481 TACGTCCACGACGCGCGGTAAATACGTAAGGTTGCAAGCGTTAATTCGGAATTAATCTGGGCGT 540
```

QY	541	AAAGGTGCGCAGGCGGCTTTGTAAGTCAGATGTGAATATCCCGGCTTAACTGGGAAAT	600
Db	541	AAAGCGTCCAGCGGCTTTTGTAAAGTCAGATGTGAATATCCCGGCTTAACTGGGAAAC	600
QY	601	TGCGTTTGAATCACTAACAAGGCTAAGAGTGTGGCAGAGGGAGGTGGAAATTCATGTGTAGAG	660
Db	601	TGCGTTTGAATCACTAACAAGGCTAAGAGTGTGGCAGAGGGAGGTGGAAATTCACGTGTAGAG	660
QY	661	TGAATATCGTAGAGATATGGAAGAACATTCGATGGCGAAGGACCTCTCGGGTTAAACACT	720
Db	661	TGAATATCGTAGAGATATGGAAGAACACCGATGGCGAAGGACGCCCTCGGGTTAAACACT	720
QY	721	GACGCTCATGACGAAACCGTGGGAGCAAAACAGATTAATATACCTCTGTATGTCCACGCC	780
Db	721	GACGCTCATGACGAAACCGTGGGAGCAAAACAGATTAATATACCTCTGTATGTCCACGCC	780
QY	781	CTAAACGATGTCACTAGTTGTTGGGCTTAATTAAGCTTGTATACGAAGCTTAACGGGTGA	840
Db	781	CTAAACGATGTCACTAGTTGTTGGGCTTAAAGCTTGTATACGAAGCTTAACGGGTGA	840
QY	841	AGTTGACCGGCTTGGGAGTAGACGTCGCGAAGATTAAACTCAAGAAATTGACGAGGACCC	900
Db	841	AGTTGACCGGCTTGGGAGTAGACGTCGCGAAGATTAAACTCAAGAAATTGACGAGGACCC	900
QY	901	GCACAAGCGGTGATATTATGTGAATTAATTCGATGCCAACGCGAAAACTTAACCTACCTT	960
Db	901	GCACAAGCGGTGATATTATGTGAATTAATTCGATGCCAACGCGAAAACTTAACCTACCTT	960
QY	961	GACATGTATGCGAATTTTCTAGAGATAGATTATGTGTC---TTGCGGAAGCGTAAACAGGTG	1017
Db	961	GACATGTATGCGAAGCCCGCTAGAGGTGGGTGTGTCGCGAAGGAGGCGTAAACAGGTG	1020
QY	1018	CTGATAGGCTGTTCGTCAGCTGTCGTGTCGAGATGTGTGGGTTAAAGTCCCGCAACGAGCGCA	1077
Db	1021	CTGATAGGCTGTTCGTCAGCTGTCGTGTCGAGATGTGTGGGTTAAAGTCTCGCAACGAGCGCA	1080
QY	1078	AACCTTGTCAATTAATTTGCCATCATTTGGTTGGGCACTTTAATGAGACTGCCGCTGACAAA	1137
Db	1081	TCCCTGTTCATTAATTTGCCATCATTCAGTTGGGCACTTTAATGAAACTGCCGCTGACAAA	1140
QY	1138	CCGAGAGAAAGTGGGGAATGACGTCAAGTCTCAATGCCCCCTTAATGGGTAAAGGCTTCAACG	1197
Db	1141	CCGAGAGAAAGTGGGGAATGACGTCAAGTCTCAATGCCCCCTTAATGGGTAAAGGCTTCAACG	1200
QY	1198	TAAATCAATGGCGCGCTACAGAGGGTTGGCAACCCGCGAGGGGAGCTAATCTCAGAAAGC	1257
Db	1201	TAAATCAATGGCGCGCTACAGAGGGTTGGCAACCCGCGAGGGGAGCTAATCTCAGAAAGC	1260
QY	1258	GCGTCGTAGTCCGGAATTCGGAAGTCTTGCAACTCGCTGTGAAGTTCGAATTCGTAGTAAT	1317
Db	1261	GCGTCGTAGTCCGGAATTCGGAAGTCTTGCAACTCGCTGTGAAGTTCGAATTCGTAGTAAT	1320
QY	1318	CGCGGATCAGCATGTGCGCGGTGAATAGTTCGCCGGGTCTTGTATACACACCGCGCTCACAC	1377
Db	1321	CGCGGATCAGCATGTGCGCGGTGAATAGTTCGCCGGGTCTTGTATACACACCGCGCTCACAC	1380
QY	1378	CATGGGAGTGGGTTTCAACAGAAAGAGGATCTAACCGTAAAGAGGCGCTTGCCACAGG	1437
Db	1381	CATGGGAGTGGGTTTCAACAGAAAGAGGATCTAACCGCAAGAGGCGCTTGCCACAGG	1440
QY	1438	TGAGATTCAATGACTGGGGTG 1457	
Db	1441	TGAGATTCAATGACTGGGGTG 1460	
RESULT 45			
AY123803 1497 bp DNA linear BCT 23-SEP-2003			
LOCUS Nitrosovibrio tenuis 16S ribosomal RNA gene, partial sequence.			
DEFINITION AY123803			
ACCESSION AY123803			
VERSION AY123803.1 GI:24474432			
KEYWORDS Nitrosovibrio tenuis			
SOURCE			

ORGANISM	REFERENCE	TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	FEATURES	ORIGIN
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	1 (bases 1 to 1497)	Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Rosier, A. and Koops, H. P.	16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal of a new lineage within the nitrosomonads	Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)			
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	2 (bases 1 to 1497)	Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roeser, A. and Wagner, M.	13330037				
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	3 (bases 1 to 1497)	Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany					
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	1. 1497	Location/Qualifiers					
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	/organism="Nitrosovibrio tenuis"						
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	/mol_type="genomic DNA"						
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	/strain="NV1"						
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	/db_xref="taxon:1233"						
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	/note="type strain"						
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	<1..>1497						
Nitrosovibrio tenuis	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.	/product="16S ribosomal RNA"						
Query Match	88.8%; Score 1294; DB 1; Length 1497;							
Best Local Similarity	93.8%; Pred. No. 2.8e-13;							
Matches 1370; Conservative	0; Mismatches 85; Indels 5; Gaps 2.							
1	ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAAACGGGCGACGATGCTTGCAT	60						
3	ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAAACGGGCGACGATGCTTGCAT	60						
61	CTGGTGGGGAATGGCGGCGGCGGAGTATGATCGAAACGTTATCCAAAGAGGGGGGTA	120						
61	CTGGTGGGGAATGGCGGCGGCGGAGTATGATCGAAACGTTATCCAAAGAGGGGGGTA	120						
121	ACGCATCGAAGATGCTGTAATACCGCATATCTTAAAGAGAAAGCAGGGGATCGAA	180						
121	ACGCATCGAAGATGCTGTAATACCGCATATCTTAAAGAGAAAGCAGGGGATCGAA	180						
121	ACGCATCGAAGATGCTGTAATACCGCATATCTTAAAGAGAAAGCAGGGGATCGAA	180						
181	GACCTTGGCGCTTTGGAGCGGCGCATGCTGATTAGCTAAGTTGGTGGGTTAAAGGCTTAC	240						
181	GACCTTGGCGCTTTGGAGCGGCGCATGCTGATTAGCTAAGTTGGTGGGTTAAAGGCTTAC	240						
181	GACCTTGGCGCTTTGGAGCGGCGCATGCTGATTAGCTAAGTTGGTGGGTTAAAGGCTTAC	240						
241	CAGGCGGAGCATGCTGTAATGCTGAGAGGACGACGACGACCTGGGACTGAGACACG	300						
241	CAGGCGGAGCATGCTGTAATGCTGAGAGGACGACGACGACCTGGGACTGAGACACG	300						
241	CAGGCGGAGCATGCTGTAATGCTGAGAGGACGACGACGACCTGGGACTGAGACACG	300						
301	GCCGAGACTCTTACCGGAGGACGACGATGGGGAATTTTGGACATGGGCGCAAGCCTGATC	360						
301	GCCGAGACTCTTACCGGAGGACGACGATGGGGAATTTTGGACATGGGCGCAAGCCTGATC	360						
301	GCCGAGACTCTTACCGGAGGACGACGATGGGGAATTTTGGACATGGGCGCAAGCCTGATC	360						
361	CAGCAATGCCGCGTGAAGAGAGGCTTGGGTTGTAAGCTCTTTCACTGAGAGAA	420						
361	CAGCAATGCCGCGTGAAGAGAGGCTTGGGTTGTAAGCTCTTTCACTGAGAGAA	420						
361	CAGCAATGCCGCGTGAAGAGAGGCTTGGGTTGTAAGCTCTTTCACTGAGAGAA	420						
421	AAAGGTTACGGTAATTAATCGTACATCAAGACCGTATGACGAGAGAGCAACCGGTAAC	480						
421	AAAGGTTACGGTAATTAATCGTACATCAAGACCGTATGACGAGAGAGCAACCGGTAAC	480						
421	AAAGGTTACGGTAATTAATCGTACATCAAGACCGTATGACGAGAGAGCAACCGGTAAC	480						
481	TACGTGCAGACGCGCGGTAATAGTAGGGGTCAAGAGGTTAATCGAAATTAAGTGGCGT	540						
481	TACGTGCAGACGCGCGGTAATAGTAGGGGTCAAGAGGTTAATCGAAATTAAGTGGCGT	540						
481	TACGTGCAGACGCGCGGTAATAGTAGGGGTCAAGAGGTTAATCGAAATTAAGTGGCGT	540						
541	AAAGGTTACGGTAATTAATCGTACATCAAGACCGTATGACGAGAGAGCAACCGGTAAC	600						
541	AAAGGTTACGGTAATTAATCGTACATCAAGACCGTATGACGAGAGAGCAACCGGTA							

```
Db 601 TCCGTTGAAAATCAAGAGCTAGAGTGTGCAAGAGGGGGGTGGAATTCACGCTGTAAGCA 660
Qy 661 TGAATATGCGTAGAGATATGGAAGAAATCATGATGCGAAGGACGCTCTGGGTTAACT 720
Db 661 TGAATATGCGTAGAGATATGGAAGAAATCATGATGCGAAGGACGCTCTGGGTTAACT 720
Qy 721 GACGCTCATGCAAGAAAGCGTGGGAGCAACAAGATTAGATACCTGCTGATCCACGCG 780
Db 721 GACGCTCATGCAAGAAAGCGTGGGAGCAACAAGATTAGATACCTGCTGATCCACGCG 780
Qy 781 CTAACAGATGTCATAGTTGTTGGGCTTATTAAGGCTTGTAACGAACTAACCGTGA 840
Db 781 CTAACAGATGTCATAGTTGTTGGGCTTATTAAGGCTTGTAACGAACTAACCGTGA 840
Qy 841 AGTTACAGCGCTGGGAGAGTACGCTGCAAGATTAATACTCAAGAAATTTGACGGGAGCC 900
Db 841 AGTTACAGCGCTGGGAGAGTACGCTGCAAGATTAATACTCAAGAAATTTGACGGGAGCC 900
Qy 901 GCAACAAGCGGTGATTAATGATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTT 960
Db 901 GCAACAAGCGGTGATTAATGATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTT 960
Qy 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTGGGAAACGCTAACAGAGTG 1017
Db 961 GACATGTAGCGAAGCCCGCTGAGAGTGGGCGTCCCGAAAGGAGCGGTAAACAGAGTG 1020
Qy 1018 CTGCATGCGTGTCTGCAAGCTGCTGTGCTGATGATGTTGGTTAATGTCGCAACGAGCG 1077
Db 1021 CTGCATGCGTGTCTGCAAGCTGCTGTGCTGATGATGTTGGTTAATGTCGCAACGAGCG 1080
Qy 1078 ACCCTGTCAATTAATGTCATATTTGTTGGGCACTTAATGAGACTGCGGGTACAA 1137
Db 1081 ACCCTGTCAATTAATGTCATATTTGTTGGGCACTTAATGAGACTGCGGGTACAA 1140
Qy 1138 CCGGAGGAAGTGGGGATGACGTCATAGTCTCATGAGCCCTTATGGGTAGGGCTTACAG 1197
Db 1141 CCGGAGGAAGTGGGGATGACGTCATAGTCTCATGAGCCCTTATGGGTAGGGCTTACAG 1200
Qy 1198 TAAATCAATGCGCGCTGACAGAGGTTGCCAACCCCGGAGGGAGGATTAATTCAGAAAG 1257
Db 1201 TAAATCAATGCGCGCTGACAGAGGTTGCCAACCCCGGAGGGAGGATTAATTCAGAAAG 1260
Qy 1258 GCGTGTAGTCCGGATCGGAGTCTGCACTCGACTCCGTAAGTGGGAATGCTATGTAAT 1317
Db 1261 GCGTGTAGTCCGGATCGGAGTCTGCACTCGACTCCGTAAGTGGGAATGCTATGTAAT 1320
Qy 1318 CCGGATCAGCATGTCGCGGTAATACGTTCCCGGGTCTTGTACACACCGCGTCAAC 1377
Db 1321 CCGGATCAGCATGTCGCGGTAATACGTTCCCGGGTCTTGTACACACCGCGCGTCAAC 1380
Qy 1378 CATGGAGTGGGTTTCAACGAAGCAGTAGTCTTAACCGTAAGAGAGGCGCTTGGCACG 1437
Db 1381 CATGGAGTGGGTTTCAACGAAGCAGTAGTCTTAACCGTAAGAGAGGCGCTTGGCACG 1440
Qy 1438 TGAGATTGATGACTGGGCTG 1457
Db 1441 TGAGATTGATGACTGGGCTG 1460

RESULT 46
AY123808 1497 bp DNA linear BCT 23-SEP-2003
LOCUS Nitrospiroplasma sp. Nsp1 16S ribosomal RNA gene, partial sequence.
DEFINITION AY123808
ACCESSION AY123808
VERSION AY123808.1 GI:24474437
KEYWORDS
SOURCE
ORGANISM Nitrospiroplasma sp. Nsp1
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrospiroplasma.
REFERENCE
1 (bases 1 to 1497)
Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roeser,A. and
```

```

TITLE      Koops,H.-P.
            16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial
            ammonia-oxidizing isolates: extension of the dataset and proposal
            of a new lineage within the nitrosomonads
JOURNAL    Inc. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)
PUBMED     13130037
REFERENCE  2 (bases 1 to 1497)
AUTHORS    Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Roeser,A. and
            Wagner,M.
TITLE      Journal
            Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische
            Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES   location/Qualifiers
            source          1..1497
                           /organism="Nitrospiroplasma sp. Nsp1"
                           /mol_type="genomic DNA"
                           /strain="Nsp1"
                           /db_xref="taxon:136547"
                           <1..>1497
                           /product="16S ribosomal RNA"
ORIGIN
Query Match      88.8%; Score 1294; DB 1; Length 1497;
Beet Local Similarity 93.8%; Pred. No. 2,8e-13;
Matches 1370; Conservative 0; Mismatches 85; Indels 5; Gaps 2;

Qy 1 ATGAAAGCGTCGCGCATGCTTTACATGCAATGCAAGTCAAGCGGACGAGTGCAT 60
Db 3 ATGAAAGCGTCGCGCATGCTTTACATGCAATGCAAGTCAAGCGGACGAGTGCAT 60
Qy 61 CTGTGCGGAGTGGCGGACGCGGTGATGATGATCATGCAACGATTCAGAAAGGGGGTGA 120
Db 61 CTGTGCGGAGTGGCGGACGCGGTGATGATGATCATGCAACGATTCAGAAAGGGGGTGA 120
Qy 121 ACGCATGAAAGATGCTTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATCGAAA 180
Db 121 ACGCATGAAAGATGCTTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATCGCAA 180
Qy 181 GACCTTGCGCTTTTGAGACGGCGCATGTCGATTAAGTATGAGTGGGTGAAAGGCTTAC 240
Db 181 GACCTTGCGCTTTTGAGACGGCGCATGTCGATTAAGTATGAGTGGGTGAAAGGCTTAC 240
Qy 241 CAAGCGCAGATCAATGATGCTGAGAGAGACGACGACACTGGGACTGAGACAG 300
Db 241 CAAGCGCAGATCAATGATGCTGAGAGAGACGACGACACTGGGACTGAGACAG 300
Qy 301 GCCCAGACTCTTACGCGGAGCGACGATGCGGAAATTTTGGACAAATGGCGCAAGCTGATC 360
Db 301 GCCCAGACTCTTACGCGGAGCGAGCGATGCGGAAATTTTGGACAAATGGCGGAAACCTGATC 360
Qy 361 CAGCAATGCGCGTGAAGAGAGGCGCTTGGGGTGTGAAGCTCTTTCAGTCGAGAA 420
Db 361 CAGCAATGCGCGTGAAGAGAGGCGCTTGGGGTGTGAAGCTCTTTCAGTCGAGAA 420
Qy 421 AAAGCTTACGCTTAATATATCGTCACTCATGACGCTATGACAGAAAGCAACCGCTTAC 480
Db 421 AAAGCTTACGCTTATATATCCGTAATCACTGACGCTATGACAGAAAGCAACCGCTTAC 480
Qy 481 TACGTGCAAGACCGCGCGTAAATACGATAGGCTGCAAGCGTTAATCGGAATTAATCTGGGCT 540
Db 481 TACGTGCAAGACCGCGCGTAAATACGATAGGCTGCAAGCGTTAATCGGAATTAATCTGGGCT 540
Qy 541 AAAGGTGCGGAGCGGCTTTGTAACTAGATGAGTAATCCCGGGGCTTAACCTGGGAAT 600
Db 541 AAAGGTGCGGAGCGGCTTTGTAACTAGATGAGTAATCCCGGGGCTTAACCTGGGAAT 600
Qy 601 TCGCTTGAATCAATCAAGGCTAGAGTGTGCAAGAGGAGTGTGAATTCATGTTAGCAG 660
Db 601 TCGCTTGAATCAATCAAGGCTAGAGTGTGCAAGAGGAGTGTGAATTCATGTTAGCAG 660
Qy 661 TGAATATGCGTAGATATGGAAGAAATCATGATGCGGAAAGGACGCTCTGGGTTAACT 720
Db 661 TGAATATGCGTAGATATGGAAGAAATCATGATGCGGAAAGGACGCTCTGGGTTAACT 720
```

```

QY 721 GAGCCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGTGATGCCAGCC 780
DB 721 GAGCCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGTGATGCCAGCC 780
QY 781 CTAAACGATGTCATAGTGTGTGCGGCTTAAACGCGCTTGTAAACGCGGATGA 840
DB 781 CTAAACGATGTCATAGTGTGTGCGGCTTAAACGCGCTTGTAAACGCGGATGA 840
QY 841 AGTTGACCCCGTGGGAGATACCGTCCGCAAGATTAAACCTCAAGAAATTGACGGGAGCC 900
DB 841 AGTTGACCCCGTGGGAGATACCGTCCGCAAGATTAAACCTCAAGAAATTGACGGGAGCC 900
QY 901 GCACAGCGGTGATATGATGATTAATTCATGACGCAACGCAAAACCTTACCTACCTT 960
DB 901 GCACAGCGGTGATATGATGATTAATTCATGACGCAACGCAAAACCTTACCTACCTT 960
QY 961 GACATGTAGCGAAATTTTCTAGAGATAGATTAGTC---TTGGGAAACCTTAACACAGGTG 1017
DB 961 GACATGTAGCGAAAGTCCGCTGAGAGGTGATGTGCCGAAAGGGAACGGGTACACAGGTG 1020
QY 1018 CTGCAATGCGTGTGTCAGCTCGTGTGATGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1077
DB 1021 CTGCAATGCGTGTGTCAGCTCGTGTGATGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1080
QY 1078 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
DB 1081 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1138 CCGGAGGAAAGTGGGAGATGACGTCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1197
DB 1141 CCGGAGGAAAGTGGGAGATGACGTCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1200
QY 1198 TAATACATAGGCGGATGACAGAGGTTGCGCAACCGCGAGGGGAGCTAATCTGAGAAAGC 1257
DB 1201 TAATACATAGGCGGATGACAGAGGTTGCGCAACCGCGAGGGGAGCTAATCTGAGAAAGC 1260
QY 1258 GCGTCGATGCTCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1317
DB 1261 GCGTCGATGCTCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1320
QY 1318 CCGGATGACGATGCTCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1377
DB 1321 CCGGATGACGATGCTCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1380
QY 1378 CATGGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1437
DB 1381 CATGGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1440
QY 1438 TGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
DB 1441 TGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1460

```

```

RESULT 47
LOCUS AY123802 1497 bp DNA linear BCT 23-SEP-2003
DEFINITION Nitrosospiira sp. Nsp2 16S ribosomal RNA gene, partial sequence.
ACCESSION AY123802
VERSION AY123802.1 GI:24474431
KEYWORDS
SOURCE
ORGANISM Nitrosospiira sp. Nsp2
Nitrosospiira sp. Nsp2
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosospiira.
REFERENCE
AUTHORS Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Rose, A. and
Koops, H.P.
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial
ammonia-oxidizing isolates: extension of the dataset and proposal
of a new lineage within the nitrosomonads
JOURNAL Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)
PUBMED 13130037

```

```

REFERENCE 2 (bases 1 to 1497)
AUTHORS Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Rose, A. and
Wagner, M.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES
source
1..1497
/organism="Nitrosospiira sp. Nsp2"
/mol_type="genomic DNA"
/strain="Nsp2"
/db_xref="taxon:136548"
<1..>1497
/product="16S ribosomal RNA"
ORIGIN
Query Match 88.7%; Score 1292.4; DB 1; Length 1497;
Best Local Similarity 93.8%; Pred. No. 36-13;
Matches 1369; Conservative 0; Mismatches 86; Indels 5; Gaps 2;
QY 1 ATTGAACGCTGCGCGGATCTTTTACATGCAAGTGCAGACGCGATGCTTGCAAT 60
DB 3 ATTGAACGCTGCGCGGATCTTTTACATGCAAGTGCAGACGCGATGCTTGCAAT 60
QY 61 CTGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 120
DB 61 CTGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 120
QY 121 AGCATGCAAGATGCTATATACGATATATCTTAAGAGGAAAGCAGGGATGCAAA 180
DB 121 AGCATGCAAGATGCTATATACGATATATCTTAAGAGGAAAGCAGGGATGCAAA 180
QY 181 GACCTTGCGCTTTTGGAGCGGCGGATGCTGATGATGATGATGATGATGATGATGATGAT 240
DB 181 GACCTTGCGCTTTTGGAGCGGCGGATGCTGATGATGATGATGATGATGATGATGATGAT 240
QY 241 CAAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 CAAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 GCCCAGATCTTACGCGGAGGAGCAGATGCGGATTTTGGCAATGCGGAGAACTTGATC 360
DB 301 GCCCAGATCTTACGCGGAGGAGCAGATGCGGATTTTGGCAATGCGGAGAACTTGATC 360
QY 361 CAGCAATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 CAGCAATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 AAGGTTAGGTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 AAGGTTAGGTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 TAGGTCGACGACGCGGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 TAGGTCGACGACGCGGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 AAGGATGCGCAGCGGCTTTTGAAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 AAGGATGCGCAGCGGCTTTTGAAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 TCGGTTTGAATCAAGGCTTAGAGTGTGCAAGGAGGATGGAATTCATGATGACAG 660
DB 601 TCGGTTTGAATCAAGGCTTAGAGTGTGCAAGGAGGATGGAATTCATGATGACAG 660
QY 661 TGAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 TGAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 GAGCCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGTGATGCCAGCC 780
DB 721 GAGCCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGTGATGCCAGCC 780
QY 781 CTAAACGATGTCATAGTGTGTGCGGCTTAAACGCGCTTGTAAACGCGGATGA 840

```



```
|||||
Db      CTAAACGATGTCACTAGTTGTCCGCTCTTAACGACTTGGTAACCGACTAACCGGTGA 840
Qy      841 AGTTGACCGCTGGGAGTAGCGGTGCAAGATTTAAACTCAAGAGATTGACGGGAGACC 900
Db      841 AGTTGACCGCTGGGAGTAGCGGTGCAAGATTTAAACTCAAGAGATTGACGGGAGACC 900
Qy      901 GCACAAAGCGGTGATTAATGTGATTAATTCGATGCAACGGCAAAAACCTTACTCTACCTT 960
Db      901 GCACAAAGCGGTGATTAATGTGATTAATTCGATGCAACGGCAAAAACCTTACTCTACCTT 960
Qy      961 GACATGTAGCGAATTTTCTAGAGATTAAGATTAGTGC--TTCCGGAACGCTTAACACAGGTG 1017
Db      961 GACATGTAGCGAACCGCCGCTGAGAGATGGGCGTCCCGGAAGGAGCGGTAAACACAGGTG 1020
Qy      1018 CTGCATGCGCTGCTGCAAGCTGCTGTCGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCA 1077
Db      1021 CTGCATGCGCTGCTGCAAGCTGCTGTCGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCA 1080
Qy      1078 ACCCTTGTCATTAATTCGATCATTTGTTGGGCACTTTAATGAGACTGCGCGGTGACAA 1137
Db      1081 ACCCTTGTCATTAATTCGATCATTTGTTGGGCACTTTAATGAGACTGCGCGGTGACAA 1140
Qy      1138 CCGGAGGAAGTGGGGATGACGTCAAGTCTCATGGCCCTTAATGGGTAGGGCTTGACACG 1197
Db      1141 CCGGAGGAAGTGGGGATGACGTCAAGTCTCATGGCCCTTAATGGGTAGGGCTTGACACG 1200
Qy      1198 TAATACATGCGCGCTGACAGAGGTTGGCCAAACCCCGAGCGGAGCTTAATCTCAGAAAGC 1257
Db      1201 TAATACATGCGCGCTGACAGAGGTTGGCCAAACCCCGAGCGGAGCTTAATCTCAGAAAGC 1260
Qy      1258 GGTGTGATGTCGGGATCGGAGTCTGCACTCGAATCCGTAAGTGGGAATGCTGTAT 1317
Db      1261 GGTGTGATGTCGGGATCGGAGTCTGCACTCGAATCCGTAAGTGGGAATGCTGTAT 1320
Qy      1318 CGCGGATGACGATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGCTACAC 1377
Db      1321 CGCGGATGACGATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGCTACAC 1380
Qy      1378 CATGGAGTGGGTTTCAACGAAGCAGTAGTCTAAACGTAAGAGAGGCGCTTGGCCACG 1437
Db      1381 CATGGAGTGGGTTTCAACGAAGCAGTAGTCTAAACGTAAGAGAGGCGCTTGGCCACG 1440
Qy      1438 TGAGATTGATGACTGGGGT 1457
Db      1441 TGAGATTGATGACTGGGGT 1460

RESULT 48
NS16SR      1474 bp      DNA      linear      BCT 23-FEB-1995
LOCUS      NS16SR
DEFINITION Nitrosospiira briensis (strain C-128) 16S ribosomal RNA (16S rRNA)
gene.
ACCESSION  U35505.1 GI:530899
VERSION     16S ribosomal RNA; ribosomal RNA; small subunit.
KEYWORDS    Nitrosospiira briensis
SOURCE      Nitrosospiira briensis
ORGANISM    Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
            Nitrosomonadaceae; Nitrosospiira.
REFERENCE   1 (bases 1 to 1474)
AUTHORS     Teebe, A., Alm, E., Regan, J.M., Toze, S., Rittmann, B.E. and Stahl, D.A.
TITLE       Evolutionary relationships among ammonia- and nitrite-oxidizing
            bacteria
JOURNAL     J. Bacteriol. 176 (21), 6623-6630 (1994)
PUBMED     7961414
COMMENT     Original source text: Nitrosospiira briensis (strain C-128) DNA.
FEATURES
            source
            location/Qualifiers
            1..1474
            /organism="Nitrosospiira briensis"
            /mol_type="genomic DNA"
            /strain="C-128"
            /db_xref="taxon:35799"
```

```
gene
1..1474
/gene="16S rRNA"
rRNA
1..1474
/gene="16S rRNA"
/product="16S ribosomal RNA"

ORIGIN

Query Match      88.7%; Score 1292.2; DB 1; Length 1474;
Best Local Similarity 93.6%; Pred. No. 3e-13;
Matches 1367; Conservative 2; Mismatches 86; Indels 5; Gaps 2;

Qy      1  ATTGAACGCTGCGGCATGCTTTACATGCAAGTCGAACGGCAGCAGATGCTTGAT 60
Db      17  ATTGAACGCTGCGGCATGCTTTACATGCAAGTCGAACGGCAGCAGATGCTTGAT 74
Qy      61  CTGGTGGAGTGGCGGAGCGGTGATATGATCTGGAACCTATTCAGAAAGGGGGT 120
Db      75  CTGGTGGAGTGGCGGAGCGGTGATATGATCTGGAACCTATTCAGAAAGGGGGT 134
Qy      121  ACGCATGGAAGATGTGCTTAATCCGCATATCTAAGAGAGAAAGCAGGGGATCGAA 180
Db      135  ACGCATGGAAGATGTGCTTAATCCGCATATCTTACGAGAAAGAGAGGGATCGAA 194
Qy      181  GACCTTGCGCTTTTGAACGGCCGATGCTGATTAAGCTAGTTGGGTAAAGCCTAC 240
Db      195  GACCTTGCGCTTTTGAACGGCCGATGCTGATTAAGCTAGTTGGGTAAAGCCTAC 254
Qy      241  CAAGCGACGATCACTAGTATGCTCTGAGAGAGACGACCGACACTGGGACTGAGACG 300
Db      255  CAAGCGCTTGATCACTAGTATGCTCTGAGAGAGACGACCGACACTGGGACTGAGACG 314
Qy      301  GCCCAGACTCCCTACCGGAGGAGCAGTGGGGAATTTTGAACATGGGCGCAAGCCTGATC 360
Db      315  GCCCAGACTCCCTACCGGAGGAGCAGTGGGGAATTTTGAACATGGGCGCAAGCCTGATC 374
Qy      361  CAGCAATCCCGCTGAGTGAAGAAAGCCCTTCGGGTGTAAGCTCTTCACTGAGAA 420
Db      375  CAGCAATCCCGCTGAGTGAAGAAAGCCCTTCGGGTGTAAGCTCTTCACTGAGAA 434
Qy      421  AAAGTTACGGTAAATTAATCTGTACTATGACCGTATGACAGAAAGACCGGCTAAC 480
Db      435  AAAGTTACGGTAAATTAATCCGTGATCACTGACCGTATGACAGAAAGACCGGCTAAC 494
Qy      481  TACGTGCGAGAGCGCGGCTAATCGTAGGGTGCAGCGCTTAATCGAATTAAGTGGGCT 540
Db      495  TACGTGCGAGAGCGCGGCTAATCGTAGGGTGCAGCGCTTAATCGAATTAAGTGGGCT 554
Qy      541  AAAGGTGCGCAGGCGGCTTTGTAACTGATGATGAATCCCGGGCTTAACTGGGAAT 600
Db      555  AAAGGTGCGCAGGCGGCTTTGTAACTGATGATGAATCCCGGGCTTAACTGGGAAT 614
Qy      601  TGCCTTTAAACTACAAAGCTAGAGTGTGCGAGAGGAGGTGGAATTCATGTTAGAG 660
Db      615  TGCCTTTAAACTACAAAGCTAGAGTGTGCGAGAGGAGGTGGAATTCATGTTAGAG 674
Qy      661  TGAATGCTGAGATATGAGAAACATGATGCGAAGGAGCGCTCTGGGTTAACTACT 720
Db      675  TGAATGCTGAGATATGAGAAACATGATGCGAAGGAGCGCTCTGGGTTAACTACT 734
Qy      721  GACGCTCATGACGAAGACGTTGGGAGCAACAGAGTTAGTACCTTGATGTCACGCC 780
Db      735  GACGCTCATGACGAAGACGTTGGGAGCAACAGAGTTAGTACCTTGATGTCACGCC 794
Qy      781  CTAAACGATGTCAACTAGTGTGGGCTTTATAGGCTTGTGAACGAAGCTAACGGCG 840
Db      795  CTAAACGATGTCAACTAGTGTGGGCTTTATAGGCTTGTGAACGAAGCTAACGGCG 854
Qy      841  AGTTGACCGCTGGGAGTACGTCGCAAGATTAAACTCAAGAGAAATTGACGGGAGACC 900
Db      855  AGTTGACCGCTGGGAGTACGTCGCAAGATTAAACTCAAGAGAAATTGACGGGAGACC 914
Qy      901  GCACAAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAATAAACCCTTAACCTT 960
```



Db 915 GCACAAGCGGTGATTTATGTGATTAATTCATGCAAGCGGAAAACTTACCTACCTT 974  
Qy 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTCCGGAACGCTTAACACAGGTG 1017  
Db 975 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTCCGGAACGCTTAACACAGGTG 1034  
Qy 1018 CTGCATGCGTGTCTGTCAGCTCTGTCTGATGATGTTGGTTAAATGCCGCAAGAGCGCA 1077  
Db 1035 CTGCATGCGTGTCTGTCAGCTCTGTCTGATGATGTTGGTTAAATGCCGCAAGAGCGCA 1094  
Qy 1078 ACCCTGTCAATTAATGTCATCATTTGGTGGCACTTTAATGAGACTGCCGCTGACAA 1137  
Db 1095 ACCCTGTCAATTAATGTCATCATTTGGTGGCACTTTAATGAGACTGCCGCTGACAA 1154  
Qy 1138 CCGGAGGAAGGTGGGGATGACGTCAGTCTCTCATGGCCCTTAATGGTTAGGGCTTCAACAG 1197  
Db 1155 CCGGAGGAAGGTGGGGATGACGTCAGTCTCTCATGGCCCTTAATGGTTAGGGCTTCAACAG 1214  
Qy 1198 TAAATCAATGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAAGC 1257  
Db 1215 TAAATCAATGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAAGC 1274  
Qy 1258 GCGTGTAGTCCGGAATCGGAGTCTGCAATCTCCCTGGAATCGGAATCGCTAGTAAT 1317  
Db 1275 GCGTGTAGTCCGGAATCGGAGTCTGCAATCTCCCTGGAATCGGAATCGCTAGTAAT 1334  
Qy 1318 CCGGATACGACGATGCGCGGTGAATAGTTCCCGGGCTTGTACACACCGCCCTGACAC 1377  
Db 1335 CCGGATACGACGATGCGCGGTGAATAGTTCCCGGGCTTGTACACACCGCCCTGACAC 1394  
Qy 1378 CATGGAGTGGGTTTTCACGAGAGAGTAGTCTTAACCGTAAAGAGGGCGCTTGCCACGG 1437  
Db 1395 CATGGAGTGGGTTTTCACGAGAGAGTAGTCTTAACCGTAAAGAGGGCGCKKNCAGG 1454  
Qy 1438 TGAGATTCACTGCTGGGGTG 1457  
Db 1455 TGAGATTCACTGCTGGGGTG 1474

RESULT 49  
AY684260 1458 bp DNA linear BCT 28-DEC-2004  
LOCUS Nitrosospira sp. GS832 16S ribosomal RNA gene, partial sequence.  
DEFINITION  
ACCESSION AY684260  
VERSION AY684260.1 GI:51012456  
KEYWORDS  
SOURCE Nitrosospira sp. GS832  
ORGANISM Nitrosospira sp. GS832  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira.  
REFERENCE  
AUTHORS 1. (bases 1 to 1458)  
TITLE Ida,T., Satoh,M., Yabe,R., Takahashi,R. and Tokuyama,T.  
JOURNAL Identification of Genus Nitrosospiro, Ammonia-Oxidizing Bacteria,  
by Comparison of N-Terminal Amino Acid Sequences of  
Phosphoglycerate Kinase  
AUTHORS J. Biosci. Bioeng. 98 (5), 380-383 (2004)  
TITLE 2. (bases 1 to 1458)  
JOURNAL Takahashi,R., Yabe,R. and Tokuyama,T.  
REFERENCE Direct Submission  
AUTHORS Submitted (12-JUL-2004) Agricultural and Biological Chemistry,  
JOURNAL College of Bioresource Sciences, Nihon University, 1866 Kameino,  
Fujisawa, Kanagawa 252-8510, Japan  
FEATURES  
Source location/Qualifiers  
1..1458  
/organism="Nitrosospira sp. GS832"  
/mol\_type="genomic DNA"  
/strain="GS832"  
/db\_xref="taxon:289297"  
<1..>1458  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 88.6%; Score 1290.8; DB 1; Length 1458;

Best Local Similarity 93.7%; Pred. No. 3-2e-13;  
Matches 1368; Conservative 0; Mismatches 87; Indels 5; Gaps 2;  
Qy 1 ATTGAACGCTGCGGCATGCTTTTACATGCAAGTCGAAACGGCAGCAGATGCTTCAT 60  
Db 1 ATTGAACGCTGCGGCATGCTTTTACATGCAAGTCGAAACGGCAGCAGATGCTTCAT 58  
Qy 61 CTGGTGGGAGTGGGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 59 CTGGTGGGAGTGGGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 118  
Qy 121 AGCATGCAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGAGGGGATCGAA 180  
Db 119 AGCATGCAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGAGGGGATCGAA 178  
Qy 181 GACCTTGCGCTTTTGGAGCGCGCATGTCTGATTAAGTAACTGATGTTGGTGGTAAAGGCTAC 240  
Db 179 GACCTTGCGCTTTTGGAGCGCGCATGTCTGATTAAGTAACTGATGTTGGTGGTAAAGGCTAC 238  
Qy 241 CAAGCGACATCAATGTTGGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
Db 239 CAAGCGTTCGATCAGTATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 298  
Qy 301 GCCGAGCTCTTACGAG 360  
Db 299 GCCGAGCTCTTACGAG 358  
Qy 361 CAGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
Db 359 CAGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418  
Qy 421 AAAGTTACGTTAATATCTGATCTCATGACGATTCGACAGAGAGAGAGAGAGAGAGAGAGAG 480  
Db 419 AAAGTTACGTTAATATCTGATCTCATGACGATTCGACAGAGAGAGAGAGAGAGAGAGAGAG 478  
Qy 481 TACGTGCCAGACCGCGGTAAATACGTAAGGTGCAAGCGTTAATCGAATTACTGGGCGT 540  
Db 479 TACGTGCCAGACCGCGGTAAATACGTAAGGTGCAAGCGTTAATCGAATTACTGGGCGT 538  
Qy 541 AAAGGTGCGCAGCGCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 539 AAAGGTGCGCAGCGCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 598  
Qy 601 TCGCTTTGAACTACAGGCTAGATGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
Db 599 TCGCTTTGAACTACAGGCTAGATGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658  
Qy 661 TGAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 659 TGAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
Qy 721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATTCACGCC 780  
Db 719 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATTCACGCC 778  
Qy 781 CTAAACGATGCTCACTAGTTGTTGGCCCTTATTAAGCTTGTGTAACGAAGTAAACGGGTGA 840  
Db 779 CTAAACGATGCTCACTAGTTGTTGGCCCTTATTAAGCTTGTGTAACGAAGTAAACGGGTGA 838  
Qy 841 AGTTGACCGCTGGGGAGTACCGTTCGCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 900  
Db 839 AGTTGACCGCTGGGGAGTACCGTTCGCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 898  
Qy 901 GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGGGAAAACTTACCTACCTT 960  
Db 899 GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGGGAAAACTTACCTACCTT 958  
Qy 961 GACATGTAGCGAATTTTCTAGATATGATTAATGTC---TTCCGGAACGCTTAACACAGGTG 1017  
Db 959 GACATGTAGCGAATTTTCTAGATATGATTAATGTC---TTCCGGAACGCTTAACACAGGTG 1018  
Qy 1018 CTGCATGCGTGTCTGTCAGCTCTGTCTGATGATGTTGGTTAAATGCCGCAAGAGCGCA 1077

Db 1019 CTGCATGGCTGTCGTACAGCTGTCGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCA 1078  
Qy 1078 ACCCTTGTCATTAAATTTGCCATCATTTGGTGGCACTTTAATGAGACTGCGGTGACAA 1137  
Db 1079 ACCCTTGTCATTAAATTTGCCATCATTTGGTGGCACTTTAATGAGACTGCGGTGACAA 1138  
Qy 1138 CCGGAGGAAGGTGGGGATGATGATGCTCAAGTCTCAATGGCCCTTANGGGTAGGGCTTCAACG 1197  
Db 1139 CCGGAGGAAGGTGGGGATGATGATGCTCAAGTCTCAATGGCCCTTANGGGTAGGGCTTCAACG 1198  
Qy 1198 TAATACAAATGCGCGCTACAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAGAAAGC 1257  
Db 1199 TAATACAAATGCGCGCTACAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAGAAAGC 1258  
Qy 1258 GCGTGTAGTCCGATCGGATCGGAGTCTGCACTTCCTCGTAAGTGGAAATCCCTAGTAAT 1317  
Db 1259 GCGTGTAGTCCGATCGGATCGGAGTCTGCACTTCCTCGTAAGTGGAAATCCCTAGTAAT 1318  
Qy 1318 CCGGATTCAGCAATGTCGCGGTGATACGTTCCTCGGTCTTGTACACACCGCGGTACAC 1377  
Db 1319 CCGGATTCAGCAATGTCGCGGTGATACGTTCCTCGGTCTTGTACACACCGCGGTACAC 1378  
Qy 1378 CATGGAGTGGGTTTCCACGAGAGCAGTAGTCTAACCCTAAGAGAGGCGCTTGCCACCG 1437  
Db 1379 CATGGAGTGGGTTTCCACGAGAGCAGTAGTCTAACCCTAAGAGAGGCGCGCTTGCCACCG 1438  
Qy 1438 TGAGATTCTAGTACTGGGGTG 1457  
Db 1439 TGAGATTCTAGTACTGGGGTG 1458

RESULT 50  
AY856377 1458 bp DNA linear BCT 15-JAN-2005  
LOCUS AY856377  
DEFINITION Nitrosospiira sp. TCH711 16S ribosomal RNA gene, partial sequence.  
ACCESSION AY856377  
VERSION AY856377.1 GI:57338940  
KEYWORDS  
SOURCE Nitrosospiira sp. TCH711  
ORGANISM Nitrosospiira sp. TCH711  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosospiira.  
REFERENCE 1 (bases 1 to 1458)  
AUTHORS Takahashi, R., Yabe, R. and Tokuyama, T.  
TITLE Direct Submission  
JOURNAL Submitted (14-DEC-2004) Agricultural and Biological Chemistry, College of Bioresource Sciences, Nihon University, 1866 Kameino, Fujisawa, Kanagawa 252-8510, Japan

FEATURES  
source location/Qualifiers  
1..1458  
/organism="Nitrosospiira sp. TCH711"  
/mol\_type="genomic DNA"  
/strain="TCH711"  
/db\_xref="taxon:308065"  
/note="similar to Nitrosospiira (Nitrosolobus) multiformis"  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 88.6%; Score 1290.8; DB 1; Length 1458;  
Best Local Similarity 93.7%; Pred. No. 3.2e-13;  
Matches 1368; Conservative 0; Mismatches 87; Indels 5; Gaps 2;

1 ATTGAACGCTGCGCGCATCTTTACATGACATGCAAGTGAACGCGCAGCAGGATCTTGCAAT 60  
Db 1 ATTGAACGCTGCGCGCATCTTTACATGACATGCAAGTGAACGCGCAGCAGGATCTTGCAAT 58  
Qy 61 CTGGTGGCAGTGGCGGAGCGGGTGAATATGATCGGAACGTAATCCAGAAAGAGGGGGGTA 120  
Db 59 CTGGTGGCAGTGGCGGAGCGGGTGAATATGATCGGAACGTAATCCAGAAAGAGGGGGGTA 118  
Qy 121 ACGCATCGAAAGATGCTAATACGATATATCTTAAGAGAGAAAGCAGGGGATCGAAA 180  
Db 121 ACGCATCGAAAGATGCTAATACGATATATCTTAAGAGAGAAAGCAGGGGATCGAAA 180

Db 119 ACGCAGCAAGGTGTGCTAATACCGCATATCTCTGAGAGAAAGAGAGGGATCCGCA 178  
Qy 181 GACCTTGCGCTTTTGGAGCGGCCGATGTCTGATTAAGTATTTGGGTGTAAGGCTTAC 240  
Db 179 GACCTTGCGCTTTTGGAGCGGCCGATGTCTGATTAAGTATTTGGGTGTAAGGCTTAC 238  
Qy 241 CAAGCGCAGTACGTAGTATGTTGCTGAGAGAGACACAGCCACTGGGACTGAGACAG 300  
Db 239 CAAGCGCAGTACGTAGTATGTTGCTGAGAGAGACACAGCCACTGGGACTGAGACAG 298  
Qy 301 GCCAGACTCTCTACGGAGGAGCAGAGTGGGAAATTTTGAACAATGCGCGCAAGCTTATC 360  
Db 299 GCCAGACTCTCTACGGAGGAGCAGAGTGGGAAATTTTGAACAATGCGCGCAAGCTTATC 358  
Qy 361 CAGCAATCCCGCTGATGAGAGAAAGGCTTTGGGTTGTAAGCTCTTTCATCCAGAAAG 420  
Db 359 CAGCAATCCCGCTGATGAGAGAAAGGCTTTGGGTTGTAAGCTCTTTCATCCAGAAAG 418  
Qy 421 AAGGTTACGGTAAATATCGTACTCATGACGGTATCGACAGAAAGACCGGCTTAC 480  
Db 419 AAGGTTACGGTAAATATCGTACTCATGACGGTATCGAGAAAGAAAGACCGGCTTAC 478  
Qy 481 TACGTCCAGCAGCGCGCGTAAATCGTAGGGTGCAGAGCTTAATCGAAATTAATCTGGGCGT 540  
Db 479 TACGTCCAGCAGCGCGCGTAAATCGTAGGGTGCAGAGCTTAATCGAAATTAATCTGGGCGT 538  
Qy 541 AAGGTTACGGTAAATATCGTACTCATGAGTGGTGAATCCCGGGGCTTAATCTGGGAAAT 600  
Db 539 AAGGTTACGGTAAATATCGTACTCATGAGTGGTGAATCCCGGGGCTTAATCTGGGAAAT 598  
Qy 601 TGCGTTGAAACTCAAGAGCTAGAGTGTGAGAGAGAGGTGAATTCATGTGTAGCAG 660  
Db 599 TGCGTTGAAACTCAAGAGCTAGAGTGTGAGAGAGAGGTGAATTCATGTGTAGCAG 658  
Qy 661 TGAATTCGTAGATATGAGAAACATCATGTGCGAAAGCAGCTCTCTGGGTTAACACT 720  
Db 659 TGAATTCGTAGATATGAGAAACATCATGTGCGAAAGCAGCTCTCTGGGTTAACACT 718  
Qy 721 GACGCTCATGCAACGAACCGTGGGAGCAACAGAAATTAGTACCTGTGTGTCCAGGCC 780  
Db 719 GACGCTCATGCAACGAACCGTGGGAGCAACAGAAATTAGTACCTGTGTGTGTCCAGGCC 778  
Qy 781 CTTAAACGATGTCAACTGATGTTGGGCTTTTATAGGCTTGTGAACGAAGCTTAACGCGTGA 840  
Db 779 CTTAAACGATGTCAACTGATGTTGGGCTTTTATAGGCTTGTGAACGAAGCTTAACGCGTGA 838  
Qy 841 AGTTGACCGCTGGGAGTACGCTGCGCAAGTAAACTCAAGAGAAATGACGCGGACCC 900  
Db 839 AGTTGACCGCTGGGAGTACGCTGCGCAAGTAAACTCAAGAGAAATGACGCGGACCC 898  
Qy 901 GCAACAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAACCTTACCTTACCTT 960  
Db 899 GCAACAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAACCTTACCTTACCTT 958  
Qy 961 GACATGTAGCAATTTTCTAGATATGATTTAGTTC---TTGCGGAGACGCTAACACAGGTG 1017  
Db 959 GACATGTAGCAATTTTCTAGATATGATTTAGTTC---TTGCGGAGACGCTAACACAGGTG 1018  
Qy 1018 CTGCATGGCTGTGCTGACGCTGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1077  
Db 1019 CTGCATGGCTGTGCTGACGCTGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1078  
Qy 1078 ACCCTTGTCATTAAATTTGCCATCATTTGGTGGCACTTTAATGAGACTGCGGTGACAA 1137  
Db 1079 ACCCTTGTCATTAAATTTGCCATCATTTGGTGGCACTTTAATGAGACTGCGGTGACAA 1138  
Qy 1138 CCGGAGGAAGGTGGGGATGATGATGCTCAAGTCTCAATGGCCCTTANGGGTAGGGCTTCAACG 1197  
Db 1139 CCGGAGGAAGGTGGGGATGATGATGCTCAAGTCTCAATGGCCCTTANGGGTAGGGCTTCAACG 1198  
Qy 1198 TAATACAAATGCGCGCTACAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAGAAAGC 1257  
Db 1199 TAATACAAATGCGCGCTACAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAGAAAGC 1258

QY 1258 GCGTCGATCCGATCGAGTCTGCACTCCGTAAGTCGGAATCGCTAGTAAT 1317  
DB 1259 GGGTCGATCCGATCGAGTCTGCACTCCGTAAGTCGGAATCGCTAGTAAT 1318  
QY 1318 CCGGATCGATCGATCGGTCGATCGTTCGCGGATCTTGTACACACCGCCCTGACAC 1377  
DB 1319 CCGGATCGATCGATCGGTCGATCGTTCGCGGATCTTGTACACACCGCCCTGACAC 1378  
QY 1378 CATTGAGTGGGTCGATCGAGAGGATGATCTAACCGTAAGAGAGGCGCTTGACACG 1437  
DB 1379 CATTGAGTGGGTCGATCGAGAGGATGATCTAACCGTAAGAGAGGCGCTTGACACG 1438  
QY 1438 TGAGATTCATGACTGGGCTG 1457  
DB 1439 TGAGATTCATGACTGGGCTG 1458

RESULT 51  
AY123790 1497 bp DNA linear BCT 23-SEP-2003  
LOCUS Nitrosospira sp. Nsp62 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY123790  
ACCESSION AY123790  
VERSION AY123790.1 GI:24474419  
KEYWORDS  
SOURCE Nitrosospira sp. Nsp62  
ORGANISM Nitrosospira sp. Nsp62  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira.  
REFERENCE  
AUTHORS 1 (bases 1 to 1497)  
Purkholz,U., Wagner,M., Timmermann,G., Pommerening-Roeser,A. and  
Koops,H.P.  
16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
JOURNAL  
PUBMED 13130037  
REFERENCE 2 (bases 1 to 1497)  
Purkholz,U., Timmermann,G., Koops,H.-P., Pommerening-Roeser,A. and  
Wagner,M.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1..1497  
/organism="Nitrosospira sp. Nsp62"  
/mol\_type="genomic DNA"  
/strain="Nsp62"  
/db\_xref="taxon:200117"  
<1..1497  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 88.6%; Score 1290.8; DB 1; Length 1497;  
Best Local Similarity 93.7%; Pred. No. 3.2e-13;  
Matches 1366; Conservative 0; Mismatches 87; Indels 5; Gaps 2;

QY 1 ATTGAAGCTGGCGGATGCTTTACATCACTCAAGCGGACGCGATGCTTGCAT 60  
DB 3 ATTGAAGCTGGCGGATGCTTTACATCACTCAAGCGGACGCGGAGGACCC-- 60  
QY 61 CTGGTGGCGATGGCGGACGGGTGATGATGATCGGACGATCGAAGAGGAGGGGTA 120  
DB 61 CTGGTGGCGATGGCGGACGGGTGATGATGATCGGACGATCGAAGAGGAGGGGTA 120  
QY 121 ACGCATGAAAGATGTGTAAATACCGCATATCTTAAGAGAGAAAGAGGGATCGAAA 180  
DB 121 ACGCATGAAAGATGTGTAAATACCGCATATCTTAAGAGAGAAAGAGGGATCGAAA 180  
QY 181 GACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTTGGTGAAGTAAAGGCTTAC 240  
DB 181 GACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTTGGTGAAGTAAAGGCTTAC 240

QY 241 CAAGCGATCGATCGATGTTGTCGAGAGGACGACCGCACCTGGGACTGAGACAG 300  
DB 241 CAAGCGATCGATCGATGTTGTCGAGAGGACGACCGCACCTGGGACTGAGACAG 300  
QY 301 GCCGACTCTCTTACCGGAGGAGGACGATGGGGAAATTTGGACAAATGGGCGCAAGCTGATC 360  
DB 301 GCCGACTCTCTTACCGGAGGAGGACGATGGGGAAATTTGGACAAATGGGCGCAAGCTGATC 360  
QY 361 CAGCAATGCCCGTCGATGGAAGAGGCTTGGGGTGTAAAGCTCTTTCAGTCGAGAAAG 420  
DB 361 CAGCAATGCCCGTCGATGGAAGAGGCTTGGGGTGTAAAGCTCTTTCAGTCGAGAAAG 420  
QY 421 AAAGTTACGGTAATTAATCGTACATGACGATGACGATGACGAGAAAGACCGGCTAAC 480  
DB 421 AAAGTTACGGTAATTAATCGTACATGACGATGACGATGACGAGAAAGACCGGCTAAC 480  
QY 481 TACGTGCGACGACCGCGGTAAATTAAGTAAAGGTCGAAAGCTTAATCGAATTACTGGCGGT 540  
DB 481 TACGTGCGACGACCGCGGTAAATTAAGTAAAGGTCGAAAGCTTAATCGAATTACTGGCGGT 540  
QY 541 AAAGGTGCGACGCGGCTTTGTAAAGTCAAGATGTAATCCCGGGCTTAACCTGGGAAT 600  
DB 541 AAAGGTGCGACGCGGCTTTGTAAAGTCAAGATGTAATCCCGGGCTTAACCTGGGAAT 600  
QY 601 TGCCTTTGAACTACAAAGCTGAGTGTGCGAGAGGGGTGGAATTCACGCTGACAG 660  
DB 601 TGCCTTTGAACTACAAAGCTGAGTGTGCGAGAGGGGTGGAATTCACGCTGACAG 660  
QY 661 TGAATGCTGAGATGATGAGAAACATCGATGCGAAGGAGCGCTCTGGGTTAACTACT 720  
DB 661 TGAATGCTGAGATGATGAGAAACATCGATGCGAAGGAGCGCGCTCTGGGTTAACTACT 720  
QY 721 GACGCTGATGACGAAAGCTGGGGAGCAAAAGCAAGATTGATACCTGATCTCCAGCGC 780  
DB 721 GACGCTGATGACGAAAGCTGGGGAGCAAAAGCAAGATTGATACCTGATCTCCAGCGC 780  
QY 781 CTTAAAGATGTAATGATGTTGGGCTTAAAGCTGATGATGATGATGATGATGATGATGAT 840  
DB 781 CTTAAAGATGTAATGATGTTGGGCTTAAAGCTGATGATGATGATGATGATGATGATGAT 840  
QY 841 AGTTGACCGCTGGGGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 841 AGTTGACCGCTGGGGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 901 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 901 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017  
DB 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017  
QY 1018 CTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077  
DB 1021 CTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
QY 1078 ACCCTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
DB 1081 ACCCTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
QY 1138 CCGGAGAGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
DB 1141 CCGGAGAGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
QY 1198 TAAATCAATGAGCGCTGACAGAGGGTTGCCAACCGGAGGGGAGCTTAATCTCAGAAAGC 1257  
DB 1201 TAAATCAATGAGCGCTGACAGAGGGTTGCCAACCGGAGGGGAGCTTAATCTCAGAAAGC 1260  
QY 1258 GCGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 1317  
DB 1261 GCGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 1320  
QY 1318 CCGGATCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 1377

Db 1321 CGCGATCGACATGTCGCGGTAAATCGTCCCGGGTCTTTGATACACCGCCCGCTCACAC 1380  
Qy 1378 CATGGAGTGGGTTTCACCAAGACAGTAGTCTAAACCGTAAGAGGGCGTTGCCACGG 1437  
Db 1381 CATGGAGTGGGTTTCACCAAGACAGTAGTCTAAACCGTAAGAGGGCGCTGCCACGG 1440  
Qy 1438 TGAGATTCTGACTGGGCTG 1457  
Db 1441 TGAGATTCTGACTGGGCTG 1460

RESULT 52  
NTL16SR  
LOCUS NTL16SR 1478 bp DNA linear BCT 23-FEB-1995  
DEFINITION Nitrosolobus multiformis (ATCC 25196) 16S ribosomal RNA (16S rRNA)  
gene.  
ACCESSION J35509  
VERSION 135509.1 GI:530900  
KEYWORDS 16S ribosomal RNA; ribosomal RNA; small subunit.  
SOURCE Nitrosospira multiformis  
ORGANISM Nitrosospira multiformis  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira.  
REFERENCE 1 (bases 1 to 1478)  
Teekle, A., Alm, E., Regan, J.M., Toze, S., Rittmann, B. R. and Stahl, D.A.  
Evolutionary relationships among ammonia- and nitrite-oxidizing  
bacteria  
J. Bacteriol. 176 (21), 6623-6630 (1994)  
JOURNAL PUBMED 7961414  
COMMENT Original source text: Nitrosolobus multiformis (library: ATCC  
25196) DNA.

FEATURES  
source  
location/Qualifiers  
1..1478  
/organism="Nitrosospira multiformis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1231"  
/cfeature\_lib="ATCC 25196"  
1..1478  
/gene="16S rRNA"  
1..1478  
/gene="16S rRNA"  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 88.5%; Score 1289.4; DB 1; Length 1478;  
Best Local Similarity 93.6%; Pred. No. 3.4e-13;  
Matches 1366; Conservative 0; Mismatches 89; Indels 5; Gaps 2;  
1 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCGAACGGCAGCAGCGATGCTGCAT 60  
Db 21 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCGAACGGCAGCAGCGGCAAC-- 78  
Qy 61 CTGTGGCGAGTGGCGGACGCGGTGAGTAATGATCGGAACGTATCCAGAAAGAGGGGAGTA 120  
Db 79 CTGTGGCGAGTGGCGGACGCGGTGAGTAATGATCGGAACGTATCCAGAAAGAGGGGAGTA 138  
Qy 121 ACGCATCGAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAA 180  
Db 139 ACGCATCGAAGATGTGCTAATACCGCATATATCTCGAGAGAAAGCAGGGGATCGCAA 198  
Qy 181 GACCTTGGGCTTTTGAAGCGCGCGATGTCGTAGTTAGCTAGTTGGTGGGTAAGGCGTAC 240  
Db 199 GACCTTGGGCTTTTGAAGCGCGCGATGTCGTAGTTAGCTAGTTGGTGGGTAAGGCGTAC 258  
Qy 241 CAAGCGACGATCAGTAATGTGCTGAGAGAGACAGCAGCCACACTGGGACTGAGACAG 300  
Db 259 CAAGCGATCGATCAGTAATGTGCTGAGAGAGACAGCAGCCACACTGGGACTGAGACAG 318  
Qy 301 GCCCAGACTCTTACCGGAGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCCTGATC 360  
Db 319 GCCCAGACTCTTACCGGAGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCCTGATC 378

Qy 361 CAGCAATGCGCGTGAAGTAAGAAAGCCTTCGCGGTTGTAAGTCTTTCACTCGAAGAA 420  
Db 379 CAGCATTGCGCGTGAAGTAAGAAAGCCTTCGCGGTTGTAAGTCTTTCACTCGAAGAA 438  
Qy 421 AAAGTTACGGTAAATATCTGTAATCTGATGACGCTATCGACAGAAAGAACCGGCTAAC 480  
Db 439 AAAGTTACGGTAAATATCTGTAATCTGATGACGCTATCGAAGAAAGAACCGGCTAAC 498  
Qy 481 TACGTGCCAGACCGCGGTAAATCTAGAGGTGCAAGCGTTAATCGAAATTACTGGGGCT 540  
Db 499 TACGTGCCAGACCGCGGTAAATCTAGAGGTGCAAGCGTTAATCGAAATTACTGGGGCT 558  
Qy 541 AAAGGTGCGCAGCGCGCTTGTAAATGATGATGTAATCCCGGGCTTAACTGGGAAAT 600  
Db 559 AAAGGTGCGCAGCGCGCTTGTAAATGATGATGTAATCCCGGGCTTAACTGGGAAAT 618  
Qy 601 TCGGTTGAAACTAACAGGCTAGAGTGGGAGAGAGGAGTGAATTCATGTTAGCAG 660  
Db 619 TCGGTTGAAACTAACAGGCTAGAGTGGGAGAGAGGAGTGAATTCATGTTAGCAG 678  
Qy 661 TGAATATGCTGAGATATGGAAGAACATCGATGGCGAAGGAGCGCTCGTGGGTTAACCT 720  
Db 679 TGAATATGCTGAGATATGGAAGAACATCGATGGCGAAGGAGCGCTCGTGGGTTAACCT 738  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGAAATTAGTACCTGTAGTCCAGCGCC 780  
Db 739 GACGCTCATGACGAAAGCGTGGGAGCAACAGAAATTAGTACCTGTAGTCCAGCGCC 798  
Qy 781 CTAAACGATGTCATAGTGTGGGCTTTATAGGCTGTGTAACGAAGCTAACGCGGTGA 840  
Db 799 CTAAACGATGTCATAGTGTGGGCTTTATAGGCTGTGTAACGAAGCTAACGCGGTGA 858  
Qy 841 AGTTGACCGGCTGGGAGTACGCTCGCAAGTTTAACTCAAGAAATTGACGGGGAGCC 900  
Db 859 AGTTGACCGGCTGGGAGTACGCTCGCAAGTTTAACTCAAGAAATTGACGGGGAGCC 918  
Qy 901 GCACAAGCGGTGAGTTATGTGATTTAATTCGATCGAACCGAATAACTTACTACCTCTT 960  
Db 919 GCACAAGCGGTGAGTTATGTGATTTAATTCGATCGAACCGAATAACTTACTACCTCTT 978  
Qy 961 GACATGTAGCGAATTTTCTAGATAGATTAATAGTGC--TTGCGGAACGCTTAAACAGGTG 1017  
Db 979 GACATGTAGCGAATTTTCTAGATAGATTAATAGTGC--TTGCGGAACGCTTAAACAGGTG 1038  
Qy 1018 CTGCATGCGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077  
Db 1039 CTGCATGCGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1098  
Qy 1078 ACCCTGTGCTTAAATTCGATCATTTGGTGGGACTTAAATGAGACTGCGGTGACAAA 1137  
Db 1099 ACCCTGTGCTTAAATTCGATCATTTGGTGGGACTTAAATGAGACTGCGGTGACAAA 1158  
Qy 1138 CCGAGAGAGGTGGGATGACGTCAAGTCTCATGCGCTTATGAGGCTTCAACAG 1197  
Db 1159 CCGAGAGAGGTGGGATGACGTCAAGTCTCATGCGCTTATGAGGCTTCAACAG 1218  
Qy 1198 TAAATACATGCGCGGTACAGAGGCTTCCAAACCGCGAGGGGGAGCTTAATCGAAGAAC 1257  
Db 1219 TAAATACATGCGCGGTACAGAGGCTTCCAAACCGCGAGGGGGAGCTTAATCGAAGAAC 1278  
Qy 1258 GCGTGTAGTCCGATCGGAATCTGCAACTCGACTCGGTGAGTGGGATGCTAGTAAT 1317  
Db 1279 GCGTGTAGTCCGATCGGAATCTGCAACTCGACTCGGTGAGTGGGATGCTAGTAAT 1338  
Qy 1318 CGCGATTCAGCATGTGCGGATGAATACGTTCCGGGCTTTGTACACACCGCGCTCACAC 1377  
Db 1339 CGCGATTCAGCATGTGCGGATGAATACGTTCCGGGCTTTGTACACACCGCGCTCACAC 1398  
Qy 1378 CATGGAGTGGGTTTCACCAAGACAGTAGTCTAAACCGTAAGAGGGCGCTTCCACAG 1437  
Db 1399 CATGGAGTGGGTTTCACCAAGACAGTAGTCTAAACCGTAAGAGGGCGCTTCCACAG 1458  
Qy 1438 TGAATTCATGACTGGGCTG 1457

Db 1459 TGAGATTCAATGACTGGGGTG 1478

|||||

RESULT 53  
AF353163 1457 bp DNA linear BCT 01-FEB-2005  
LOCUS Nitrosospiira sp. PJA1 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF353163  
ACCESSION AF353163.1 GI:13487772

KEYWORDS  
SOURCE Nitrosospiira sp. PJA1  
ORGANISM Nitrosospiira sp. PJA1  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospiira.  
1 (bases 1 to 1457)  
Takashi, R., Sato, K. and Tokuyama, T.  
Unpublished  
2 (bases 1 to 1457)  
Takashi, R., Sato, K. and Tokuyama, T.  
Unpublished  
REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (26-FEB-2001) Agricultural and Biological Chemistry,  
Journal College of Bioresource Sciences, Nihon University, 1866 Kameino,  
Fujisawa 252-8510, Japan  
Location/Qualifiers  
1. 1457  
/organism="Nitrosospiira sp. PJA1"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:155921"  
/note="similar to Nitrosospiira (Nitrosolobus) multiformis;  
similar to Nitrosospiira (Nitrosolobus) multiformis"  
<1. >1457  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 88.5%; Score 1289.2; DB 1; Length 1457;  
Best Local Similarity 93.6%; Pred. No. 3,4e-13;  
Matches 1366; Conservative 0; Mismatches 89; Indels 4; Gaps 2;

1 ATTGAACGCTGGCGGCGATCTTTACATCAATGCAAGCGGAGCGAGCTTTCAT 60  
1 ATTGAACGCTGGCGGCGATCTTTACATCAATGCAAGCGGAGCGAGCTTTCAT 58  
61 CTGGTGGCGATGGCGGCGGATGATGATCGGAGCGATCCAGAGAGGGGGGATA 120  
59 CTGGTGGCGATGGCGGCGGATGATGATCGGAGCGATCCAGAGAGGGGGGATA 118  
121 ACCGATCGAAGATGTGCTAATACCGCATATCTTAAGAGGAAAGCAGGGGATCGAA 180  
119 ACCGATCGAAGATGTGCTAATACCGCATATCTTAAGAGGAAAGCAGGGGATCGAA 178  
181 GACCTTGGCTTTTGAAGCGCGCGATGCTGATTAAGCTTGGTGGGTTAAAGGCTTAC 240  
179 GACCTTGGCTTTTGAAGCGCGCGATGCTGATTAAGCTTGGTGGGTTAAAGGCTTAC 238  
241 CAAGGCGACGATGATGTTGTGCTGAGAGGAGCGACGACGACCTGAGACG 300  
239 CAAGGCGCTGATGATGTTGTGCTGAGAGGAGCGACGACGACCTGAGACG 298  
301 GCCGAGACTCTTAACGAGGAGCGAGGATTTTGAACAATGGGCGCAAGCTGATC 360  
299 GCCGAGACTCTTAACGAGGAGCGAGGATTTTGAACAATGGGCGCAAGCTGATC 358  
361 CAGCAATGCCGCTGATGATGAAGAGGCTTGGGTTGAAGCTTTTCACTGAGAGAA 420  
359 CAGCAATGCCGCTGATGATGAAGAGGCTTGGGTTGAAGCTTTTCACTGAGAGAA 418  
421 AAAGTTACGGTAAATATGTAATGATGATGATGATGATGATGATGATGATGATGAT 480  
419 AAAGTTACGGTAAATATGTAATGATGATGATGATGATGATGATGATGATGATGAT 478  
481 TACGTGCGACGAGCGCGGTAAATGTAAGGTGCAAGCGTTAATCGAATTAATTGGGCGT 540

Db 479 TACGTGCGACGAGCGCGGTAAATGTAAGGTGCAAGCGTTAATCGAATTAATTGGGCGT 538

|||||

RESULT 54  
AF386756 1486 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosospiira sp. R3c5 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386756  
ACCESSION AF386756

479 TACGTGCGACGAGCGCGGTAAATGTAAGGTGCAAGCGTTAATCGAATTAATTGGGCGT 538  
541 AAAGGTGCGACGAGCGCGGTAAATGTAAGGTGCAAGCGTTAATCGAATTAATTGGGCGT 600  
539 AAAGGTGCGACGAGCGCGGTAAATGTAAGGTGCAAGCGTTAATCGAATTAATTGGGCGT 598  
601 TCGGTTTGAATCAATGAGCTTGAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
599 TCGGTTTGAATCAATGAGCTTGAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 658  
661 TGAATGCGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
659 TGAATGCGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
721 GACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
719 GACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 778  
781 CTAAACGATGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
779 CTAAACGATGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838  
841 AGTTGACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
839 AGTTGACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 898  
901 GCACAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
899 GCACAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958  
961 GACATGT--AGCGAATTTTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018  
959 GACATGTGCGAGACCGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1018  
1019 TGCATGCGCTGGTGCAGCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078  
1019 TGCATGCGCTGGTGCAGCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078  
1079 CCCTTGTCAATTAATGTCATCAATTTGGTGGGACCTTAATGAGACTGCGCGTGACAAAC 1138  
1079 CCCTTGTCAATTAATGTCATCAATTTGGTGGGACCTTAATGAGACTGCGCGTGACAAAC 1138  
1139 CGAGGAGAGGTGGGAGTGAAGCTCAATGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1198  
1139 CGAGGAGAGGTGGGAGTGAAGCTCAATGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1198  
1199 AATACAAATGGCGGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1258  
1199 AATACAAATGGCGGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1258  
1259 CGTGTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1318  
1259 CGTGTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1318  
1319 GCGGATCAAGATGCTCGCGGTGAATCGTTCCCGGGTCTTGTACACACCGCCGCTCAAC 1378  
1319 GCGGATCAAGATGCTCGCGGTGAATCGTTCCCGGGTCTTGTACACACCGCCGCTCAAC 1378  
1379 ATGGAGTGGGTTTCAACGAAGACAGTATGCTTAACCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1438  
1379 ATGGAGTGGGTTTCAACGAAGACAGTATGCTTAACCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1438  
1439 GAGATTCAATGATGGGGTG 1457  
1439 GAGATTCAATGATGGGGTG 1457

VERSION AF386756.1 GI:17864833  
KEYWORDS ENV.  
SOURCE Nitrosospira sp. R3c5  
ORGANISM Nitrosospira sp. R3c5  
REFERENCE Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira; environmental samples.  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 1172936  
REFERENCE 2 (bases 1 to 1486)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condon Dr, Moorpark, CA 93021, USA  
FEATURES  
source location/Qualifiers  
1. 1486  
/organism="Nitrosospira sp. R3c5"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichment"  
/db\_xref="taxon:180550"  
/clone="R3c5"  
/environmental\_sample  
1. 1486  
/product="16S ribosomal RNA"  
ORIGIN  
trna  
Query Match 88.5%; Score 1289.2; DB 3; Length 1486;  
Beet Local Similarity 93.6%; Pred. No. 3.4e-13;  
Matches 1367; Conservative 0; Mismatches 88; Indels 5; Gaps 2;  
QY 1 ATTGAACGCTGGCGGATGCTTTACATGACGAAGTGAACGGCAGCAGCGATCTTGCAT 60  
DB 15 ATTGAACGCTGGCGGATGCTTTACATGACGAAGTGAACGGCAGCAGCGATCTTGCAT 72  
QY 61 CTGGTGGCGATGGCGGAGCGGATGATATGATGATGATGATGATGATGATGATGATGAT 120  
DB 73 CTGGTGGCGATGGCGGAGCGGATGATATGATGATGATGATGATGATGATGATGATGAT 132  
QY 121 ACGCATCGAAAGATGCTATACCGCATATCTCTAAGAGAAAGCAGGAGATCGAAA 180  
DB 133 ACGCATCGAAAGATGCTATACCGCATATCTCTAAGAGAAAGCAGGAGATCGAAA 192  
QY 181 GACCTTGGCGCTTTTGGAGCGCGCATGCTGATTAAGCTAATTGATGGGTTAAAGCCTAC 240  
DB 193 GACCTTGGCGCTTTTGGAGCGCGCATGCTGATTAAGCTAATTGATGGGTTAAAGCCTAC 252  
QY 241 CAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 253 CAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 312  
QY 301 GCGCGAGCTCTTACCGGAGGAGCAGTGGGGAATTTTGGCAATGGGCGCAAGCTGATC 360  
DB 313 GCGCGAGCTCTTACCGGAGGAGCAGTGGGGAATTTTGGCAATGGGCGCAAGCTGATC 372  
QY 361 CAGCATATGCGCGTGAAGTGAAGAGCCCTTCGGGTGTTAAAGCTCTTTTCAGTCGAGAGA 420  
DB 373 CAGCATATGCGCGTGAAGTGAAGAGCCCTTCGGGTGTTAAAGCTCTTTTCAGTCGAGAGA 432  
QY 421 AAAGGTACGGTAAATTAATCGTACTCATGACGGTATGACAGAAAGACACCGGCTAAC 480  
DB 433 AAAGGTACGGTAAATTAATCGTACTCATGACGGTATGACAGAAAGACACCGGCTAAC 492  
QY 481 TACGTGCAAGACCGCGGCTAATAGTAGGTTGCAAGCTTATCGAATTACTGAGGCT 540  
DB 493 TACGTGCAAGACCGCGGCTAATAGTAGGTTGCAAGCTTATCGAATTACTGAGGCT 552  
QY 541 AAAGGTGCGCAGCGGCTTTTGAAGTCAAGTGTGAATCCCGGGCTTAACTGAGAT 600  
DB 553 AAAGGTGCGCAGCGGCTTTTGAAGTCAAGTGTGAATCCCGGGCTTAACTGAGAT 612

QY 601 TCGCTTGAATCAAGGCTAGAGTGTGGCAGAGGAGGAGTGAATTCATGTTAGCAG 660  
DB 613 TCGCTTGAATCAAGGCTAGAGTGTGGCAGAGGAGGAGTGAATTCATGTTAGCAG 672  
QY 661 TGAATGCGTGAAGTATGAAAGACATGATGATGATGATGATGATGATGATGATGATGAT 720  
DB 673 TGAATGCGTGAAGTATGAAAGACATGATGATGATGATGATGATGATGATGATGATGAT 732  
QY 721 GACGCTATGACGAAACGTTGGGAGAGCAACAGATTAGTACCTGTTAGTCCAGCC 780  
DB 733 GACGCTATGACGAAACGTTGGGAGAGCAACAGATTAGTACCTGTTAGTCCAGCC 792  
QY 781 CTAAACGATGTCATGATGTTGTTGGCTTTTATGAGCTGTTGTTAAGCAAGCTAAGCCGTA 840  
DB 793 CTAAACGATGTCATGATGTTGTTGGCTTTTATGAGCTGTTGTTAAGCAAGCTAAGCCGTA 852  
QY 841 AGTTGACCGCTGGGAGATACGCTGCGCAAGATTAACTCAAGGAATGACGGGAGCC 900  
DB 853 AGTTGACCGCTGGGAGATACGCTGCGCAAGATTAACTCAAGGAATGACGGGAGCC 912  
QY 901 GCACAGCGGTGATTAATGATGATTAATTCATGACGACGAAACCTTAACCTT 960  
DB 913 GCACAGCGGTGATTAATGATGATTAATTCATGACGACGAAACCTTAACCTT 972  
QY 961 GACATGTAGCAATTTTCTAGATAGATTAGTGC---TTGGGAAAGCTTAACAGGTG 1017  
DB 973 GACATGTAGCAATTTTCTAGATAGATTAGTGC---TTGGGAAAGCTTAACAGGTG 1032  
QY 1018 CTGCATGCTGCTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
DB 1033 CTGCATGCTGCTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092  
QY 1078 ACCCTTGTCAATTAATGCTATCAATTTGTTGGGCACTTAAATGAGCTGCGGTGACAA 1137  
DB 1093 ACCCTTGTCAATTAATGCTATCAATTTGTTGGGCACTTAAATGAGCTGCGGTGACAA 1152  
QY 1138 CCGAGGAAAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
DB 1153 CCGAGGAAAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212  
QY 1198 TAAATCAATGCGGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAAGAAAGC 1257  
DB 1213 TAAATCAATGCGGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAAGAAAGC 1272  
QY 1258 GCGTGTATGTCGGATGCGGATCTGCACTGATCTCGGTGAATGCGGAATGCTAAT 1317  
DB 1273 GCGTGTATGTCGGATGCGGATCTGCACTGATCTCGGTGAATGCGGAATGCTAAT 1332  
QY 1318 GCGGATACGATGTCGGGTTGAAATGATGTCGGGTTGTTGTAACAACCGCGGTCAAC 1377  
DB 1333 GCGGATACGATGTCGGGTTGAAATGATGTCGGGTTGTTGTAACAACCGCGGTCAAC 1392  
QY 1378 CATGGAAGTGGGTTTCAACAGAGCAGGTAGTCTTAACCGTAAGAGAGGCGCTTGCAACG 1437  
DB 1393 CATGGAAGTGGGTTTCAACAGAGCAGGTAGTCTTAACCGTAAGAGAGGCGCTTGCAACG 1452  
QY 1438 TGAATTCATGATGCGGTTG 1457  
DB 1453 TGAATTCATGATGCGGTTG 1472  
RESULT 55  
AT123787  
LOCUS AY123787 1497 bp DNA linear BCT 23-SEP-2003  
DEFINITION Nitrosospira sp. Nsp40 16S ribosomal RNA gene, partial sequence.  
ACCESSION AY123787  
VERSION AY123787.1 GI:24474416  
KEYWORDS  
SOURCE  
ORGANISM Nitrosospira sp. Nsp40  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira.

REFERENCE 1 (bases 1 to 1497)  
AUTHORS Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roser,A. and  
Koops,H.P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
JOURNAL Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
PUBMED 13130037  
REFERENCE 2 (bases 1 to 1497)  
AUTHORS Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Roeser,A. and  
Wagner,M.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
LOCATION/Qualifiers  
FEATURES  
Source 1.1497  
/organism="Nitrosospira sp. Nep10"  
/mol\_type="genomic DNA"  
/strain="Nep10"  
/db\_xref="taxon:200114"  
<1..>1497  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 88.5%; Score 1289.2; DB 1; Length 1497;  
Best Local Similarity 93.6%; Pred. No. 3.3e-13;  
Matches 1367; Conservative 0; Mismatches 88; Indels 5; Gaps 2;  
QY 1 ATTGAACGCTGGCGGCGATGCTTTACATACATCGAAGTCGAAGCGACGACGCGATGTCAT 60  
DB 3 ATTGAACGCTGGCGGCGATGCTTTACATACATCGAAGTCGAAGCGACGACGCGGAGMAC-- 60  
QY 61 CTGGTGGCGATGGCGGCGAGCGGTGATGATGATCGAAGCGATCCAGAAGAGGGGGGTA 120  
DB 61 CTGGTGGCGATGGCGGCGAGCGGTGATGATGATCGAAGCGATCCAGAAGAGGGGGGTA 120  
QY 121 ACCGATCGAAGATGTGCTAATACCGCATATCTTAAGAGGAAGACAGGGGATCGAAA 180  
DB 121 ACCGATCGAAGATGTGCTAATACCGCATATCTTAAGAGGAAGACAGGGGATCGCAA 180  
QY 181 GACCTTCGCTTTTGAAGCGGCGCATGTCTGATTAGTTGATGAGGTAAAGGCTTAC 240  
DB 181 GACCTTCGCTTTTGAAGCGGCGCATGTCTGATTAGTTGATGAGGTAAAGGCTTAC 240  
QY 241 CAAGGGACGATCGATGTTGCTGAGAGGACGACGACCACTGGGATCTGAGACAG 300  
DB 241 CAAGGGATCGATGTTGCTGAGAGGACGACGACCACTGGGATCTGAGACAG 300  
QY 301 GCCGAGATCTCTTACGAGGAGGACGATGGGAAATTTTGAACAATGGGCGCAAGCTGATC 360  
DB 301 GCCGAGATCTCTTACGAGGAGGACGATGGGAAATTTTGAACAATGGGCGCAAGCTGATC 360  
QY 361 CAGCAATGCCGCTGATGAAAGAGCTTTGGGTTGTAAGGCTTTTCACTCGAAGAA 420  
DB 361 CAGCAATGCCGCTGATGAAAGAGCTTTGGGTTGTAAGGCTTTTCACTCGAAGAA 420  
QY 421 AAAGGTTACGGTAAATATGTACTGATGATGATGATGATGATGATGATGATGATGATGAT 480  
DB 421 AAAGGTTACGGTAAATATGTACTGATGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 481 TACGTGCAGACGCGCGGTAAATGTAAGGTGCAAGGCTTAATCGAATTACTGGGCGT 540  
DB 481 TACGTGCAGACGCGCGGTAAATGTAAGGTGCAAGGCTTAATCGAATTACTGGGCGT 540  
QY 541 AAAGGTTCCGAGCGGCTTTGTAAGTGAATGTAATGTAATGTAATGTAATGTAATGTAAT 600  
DB 541 AAAGGTTCCGAGCGGCTTTGTAAGTGAATGTAATGTAATGTAATGTAATGTAATGTAAT 600  
QY 601 TGGCTTTAAACTCAAGGCTAGTGAGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
DB 601 TGGCTTTAAACTCAAGGCTAGTGAGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
QY 661 TGAATGCGTAGATATGAAAGACATGATGCGGAAGGACCTCTCTGGGTTAACT 720

DB 661 TGAATGCGTAGATATGAAAGACATGATGCGGAAGGACCTCTCTGGGTTAACT 720  
QY 721 GACGCTCATGACGAAGAGGTGGAGCAACAGATTAATGATACCTGTAGTCCAGCGC 780  
DB 721 GACGCTCATGACGAAGAGGTGGAGCAACAGATTAATGATACCTGTAGTCCAGCGC 780  
QY 781 CTAAACGATGTCATGATGTTGTTGGGCTTTATTAAGCTTGTAAGCACTAACGGGTGA 840  
DB 781 CTAAACGATGTCATGATGTTGTTGGGCTTTATTAAGCTTGTAAGCACTAACGGGTGA 840  
QY 841 AGTTGACCGCTGGGGAGTACGTCGCAAGATTTAAACTCAAGAAATTGACGGGAGCC 900  
DB 841 AGTTGACCGCTGGGGAGTACGTCGCAAGATTTAAACTCAAGAAATTGACGGGAGCC 900  
QY 901 GCACAAAGCGGTGATTAATGATTAATTCATGATGCAACCGCAAAAACCTTACTACCTT 960  
DB 901 GCACAAAGCGGTGATTAATGATTAATTCATGATGCAACCGCAAAAACCTTACTACCTT 960  
QY 961 GACATGTACGAAATTTTCTAGATGATTAAGTGC--TTGGGAAAGCTTAACACAGGTG 1017  
DB 961 GACATGTACGAAATTTTCTAGATGATTAAGTGC--TTGGGAAAGCTTAACACAGGTG 1017  
QY 1018 CTGCATGCTGTCGTCAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAACGAGCGCA 1077  
DB 1021 CTGCATGCTGTCGTCAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAACGAGCGCA 1080  
QY 1078 ACCCTGTGATTAATGTCATCATCTTTGTTGGCACTTTTAATGACACTGCGGTGACAA 1137  
DB 1081 ACCCTGTGATTAATGTCATCATCTTTGTTGGCACTTTTAATGACACTGCGGTGACAA 1140  
QY 1138 CCGGAGGAAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
DB 1141 CCGGAGGAAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
QY 1198 TAATTAACAATGGCGCTGACAGAGGTTGCCAACCCGAGGGGAGGCTTAATCTGAAAGC 1257  
DB 1201 TAATTAACAATGGCGCTGACAGAGGTTGCCAACCCGAGGGGAGGCTTAATCTGAAAGC 1260  
QY 1258 GCGTGTGATGTCGCGATCGAGTCTGAATCTGATCTCGTGAATCGGATCGTATGAT 1317  
DB 1261 GCGTGTGATGTCGCGATCGAGTCTGGAATCTGATCTCGTGAATCGGATCGTATGAT 1320  
QY 1318 CCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTAACAACCGCCGTCAC 1377  
DB 1321 CCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTAACAACCGCCGTCAC 1380  
QY 1378 CATGGAGTGGGTTTACCAAGAGCAGGTAGTCTAAACCGTAAGAGGCGCTTGCCACGG 1437  
DB 1381 CATGGAGTGGGTTTCCACCAAGAGCAGGTAGTCTAAACCGTAAGAGGCGCTTGCCACGG 1440  
QY 1438 TGAGATTTCATGACTGGGGTG 1457  
DB 1441 TGAGATTTCATGACTGGGGTG 1460  
RESULT 56  
AY123804 1497 bp DNA linear BCT 23-SEP-2003  
LOCUS Nitrosospira sp. Nep17 16S ribosomal RNA gene, partial sequence.  
DEFINITION Nitrosospira sp. Nep17 16S ribosomal RNA gene, partial sequence.  
ACCESSION AY123804  
VERSION AY123804.1 GI:24474433  
KEYWORDS  
SOURCE  
ORGANISM  
Nitrosospira sp. Nep17  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira.  
REFERENCE  
1 (bases 1 to 1497)  
AUTHORS Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roser,A. and  
Koops,H.P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads



JOURNAL	PUBLISHED	REFERENCES	AUTHORS	TITLE	JOURNAL
Inf. J. Syst. Evol. Microbiol.	53 (Pt 5), 1485-1494 (2003)				
1313J0037	2 (bases 1 to 1497)				
Pukhoid, U., Timmermann, G., Koops, H.-P., Pommerening-Roeser, A. and Wagner, M.					
Submitted (19-jun-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany					
Location/Qualifiers					
1..1497					
/organism="Mycobacterium sp. Nsp17"					
/mol_type="genomic DNA"					
/strain="Nsp17"					
/db_xref="taxon:136546"					
<1..>.1497					
/product="16S ribosomal RNA"					
ORIGIN					
Query Match	88.5%; Score 1289.2; DB 1; Length 1497;				
Best Local Similarity	93.6%; Pred. No. 3.3e-13;				
Matches 1367; Conservative	0; Mismatches 88; Indels 5; Gaps 2				
OY	1	ATTGAACGCTGGCGCATGCTTTTACACATGCAAGTCGAAAGCAGGACGAGATGCTGCAT	60		
Dd	3	ATTGAACGCTGGCGCATGCTTTTACACATGCAAGTCGAAAGCAGGACGAGATGCTGCAT	60		
OY	61	CTGTGGCGAAGTGCGGAGCGAGTGAATAATGCAATCGAAAGTAACCAAGAAGGGGGGTA	120		
Dd	61	CTGTGGCGAAGTGCGGAGCGAGTGAATAATGCAATCGAAAGTAACCAAGAAGGGGGGTA	120		
OY	121	ACGCATCGAAGATGCTGAATACCGCATATCTTTAGAGAGAAAGCAGGGGATCGAAA	180		
Dd	121	ACGCATCGAAGATGCTGAATACCGCATATCTTTAGAGAGAAAGCAGGGGATCGCAA	180		
OY	181	GACCTTGGGCTTTTGAAGCGCCGATGCTGAATGCTAGTTGGTGGGCTTAAAGGCTTAC	240		
Dd	181	GACCTTGGGCTTTTGAAGCGCCGATGCTGAATGCTAGTTGGTGGGCTTAAAGGCTTAC	240		
OY	241	CAAGCGCAGATCACTAGTTGGTCTGAGAGAGACGACCAACCACCTGGACTGAGACAG	300		
Dd	241	CAAGCGCAGATCACTAGTTGGTCTGAGAGAGACGACCAACCACCTGGACTGAGACAG	300		
OY	301	GCCCAACTCTTAACGGGAGGCGAGTGGGAAATTTGACAATGGGCGCAAGCTGATC	360		
Dd	301	GCCCAACTCTTAACGGGAGGCGAGTGGGAAATTTGACAATGGGCGCAAGCTGATC	360		
OY	361	CAGCAATGCCGGTAGAGAAAGAGGCTTGGGTTGTAAAGCTCTTTAGTCGAGAAAC	420		
Dd	361	CAGCAATGCCGGTAGAGAAAGAGGCTTGGGTTGTAAAGCTCTTTAGTCGAGAAAC	420		
OY	421	AAAGGTTACGGTAAATTAATCTGATCACTAGACGGTATCGACAGAAAGAACCCGCTAAC	480		
Dd	421	AAAGGTTACGGTAAATTAATCTGATCACTAGACGGTATCGACAGAAAGAACCCGCTAAC	480		
OY	481	TACGTGCAGACGCGCGGTAAATCGTAGGTTGCAAGCGTTAATGGAATTAATCTGGCGT	540		
Dd	481	TACGTGCAGACGCGCGGTAAATCGTAGGTTGCAAGCGTTAATGGAATTAATCTGGCGT	540		
OY	541	AAAGGTTACGGTAAATTAATCTGATCACTAGACGGTATCGACAGAAAGAACCCGCTAAC	600		
Dd	541	AAAGGTTACGGTAAATTAATCTGATCACTAGACGGTATCGACAGAAAGAACCCGCTAAC	600		
OY	601	TGCGTTTAAACTAACAAGGCTAGATGTCGAGAGGAGGAGTGGAAATTCATGTGTAGAC	660		
Dd	601	TGCGTTTAAACTAACAAGGCTAGATGTCGAGAGGAGGAGTGGAAATTCATGTGTAGAC	660		
OY	661	TGAATGCGTAGAATGTGAAGAACACCGATGGGAAGGAGCCCCCTGGGTAAACAT	720		
Dd	661	TGAATGCGTAGAATGTGAAGAACACCGATGGGAAGGAGCCCCCTGGGTAAACAT	720		
OY	721	GACGCTCATGCAAGAAAGCTGGGAGCAACAGATTTAGATACCTGGTATGCCAGCC	780		
Dd	721	GACGCTCATGCAAGAAAGCTGGGAGCAACAGATTTAGATACCTGGTATGCCAGCC	780		

QY		761	CTAAACGATGCCA	CTACTGTGTGGGCGCTTATTAGGCTTTGGTAACGAAGCTAACGCCTGA	840
Db		761	CTAAACGATGCCA	CTACTGTGTGGGCTTATTAGGCTTTGGTAACGAAGCTAACGCCTGA	840
QY		841	AGTTACCAGCCG	CTGGGAGTAGCGTGCGCAAGATTAAA	900
Db		841	AGTTACCAGCCG	CTGGGAGTAGCGTGCGCAAGATTAAA	900
QY		901	GACAAGAAGCGG	TGATATGTGGATTATTCGATGCAAGCGCAAAACCTTACCTT	960
Db		901	GACAAGAAGCGG	TGATATGTGGATTATTCGATGCAAGCGCAAAACCTTACCTT	960
QY		961	GACATGTACCGAA	TTTTCTTAGAGATAGATTATGTGCTC---TTGGGGAAAGCCTTAACACAGGTG	101
Db		961	GACATGTACCGAA	TTTTCTTAGAGATAGATTATGTGCTC---TTGGGGAAAGCCTTAACACAGGTG	102
QY		1018	CTGCATGAGCTGT	CGTCAAGCTCGTGTCTGTGAGATGTTGGATTATAGTCCGCAACGAGCGCA	1077
Db		1021	CTGCATGAGCTGT	CGTCAAGCTCGTCAAGCTCGTGTCTGTGAGATGTTGGATTATAGTCCGCAACGAGCGCA	108
QY		1078	AACCTTTGATTTA	ATTGCGCATCATTTTGGTTGGGCACTTTAATGAGACTGTCCGGTGA	1137
Db		1081	AACCTTTGATTTA	ATTGCGCATCATTTTGGTTGGGCACTTTAATGAGACTGTCCGGTGA	1140
QY		1138	CCGAGGAAGG	TGGGAGTAGACGTCAAGCTCATGAGCCCTTATGAGGTTAGGCTTACACAG	1197
Db		1141	CCGAGGAAGG	TGGGAGTAGACGTCAAGCTCATGAGCCCTTATGAGGTTAGGCTTACACAG	1200
QY		1198	TAAATCAAATGG	CGCGCTCACAGAGGGTTGGCCAACCGCGAAGGGAGACTAATCTCAGAAAGC	1257
Db		1201	TAAATCAAATGG	CGCGCTCACAGAGGGTTGGCCAACCGCGAAGGGAGACTAATCTCAGAAAGC	1260
QY		1258	GCGTGTATGTCC	GGATTCGGAAGCTTTCGCACTTGATCTCCGTGAATGCGTAATGCTTAAT	1317
Db		1261	GCGTGTATGTCC	GGATTCGGAAGCTTTCGCACTTGATCTCCGTGAATGCGTAATGCTTAAT	1320
QY		1318	CGCGGATCAGCA	TGTCGCGGTGAATACTTCCCAGGCTTTGTACACACCGGCCCTGCACAC	1377
Db		1321	CGCGGATCAGCA	TGTCGCGGTGAATACTTCCCAGGCTTTGTACACACCGGCCCTGCACAC	1380
QY		1378	CATGGAGATGG	AGTTTCACCAAGACGAGTAGCTTAACCGTAAAGAGGGCGCTTTCACAGG	1437
Db		1381	CATGGAGATGG	AGTTTCACCAAGACGAGTAGCTTAACCGGTAAGAGAGGGCGCTTTCACAGG	1440
QY		1438	TGAGATTGATCA	CTCGGGGTG	1457
Db		1441	TGAGATTGATCA	CTCGGGGTG	1460
RESULT 57					
LOCUS	AY123805		1497 bp	DNA linear	BCT 23-SEP-2003
DEFINITION	Nitrosospiira sp. Nv6	16S	rribosomal RNA gene,	partial sequence.	
ACCESSION	AY123805				
VERSION	AY123805.1	GI:24474434			
KEYWORDS	Nitrosospiira sp. Nv6				
SOURCE	Nitrosospiira sp. Nv6				
ORGANISM	Bacteria; Proteobacteria; Betaproteobacteriales; Nitrosomonadales;				
	Nitrosomonadaceae; Nitrosospiira.				
REFERENCE	1 (bases 1 to 1497)				
AUTHORS	Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roseer,A. and Koops,H.P.				
TITLE	16S rRNA and amoa-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal of a new lineage within the nitrosomonade				
JOURNAL	Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)				
PUBMED	13130037				
REFERENCE	2 (bases 1 to 1497)				
AUTHORS	Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Roseer,A. and Wagner,M.				
TITLE	Direct Submission				



**JOURNAL** Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universität Muenchen, Am Hochanger 4, Freising 85350, Germany  
**FEATURES** location/Qualifiers  
**SOURCE** 1. .1497

FEATURES	Location/Qualifiers
source	1. .1497

**Source**

1. .1497

1. .1497

1. .1497

1. .1497

1. .1497

1. .1497

1. .1497

rRNA

Query Match	88.5%	Score 1289.2	DB 1	Length 1497
-------------	-------	--------------	------	-------------

Best Local Similarity 93.6%; Pred. No. 3.3e-13;  
Matches 1367; Conservative 0; Mismatches 88; Indels 5; Gaps 2

Matches 1367; Conservative 0; Mismatches 88; Indels 5; Gaps 2;

OY	1	ATTGAAACCTGCGGCATGCTTTACATATGCAGATCGAAGCGCACATCGGAATGCTTGTCAAT	80
Db	3	ATTGAAGCTGCGGCATGCTTTACATATGCAGATCGAAGCGCACATCGGAAGCTTGTCAAT	60
OY	61	CTGGTGGCGAGTGGCGGACCGGTAGTAAATGCATTCGGAAACGTATCCAGAAAGGGGGGATA	120
Db	61	CTGGTGGCGAGTGGCGGAAACCGCTAGTATGCATCGGAACGTATCTTCAAGTGGGGGATA	120
OY	121	AACGATCGAAAGATGTGCTAAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAA	180
Db	121	ACGCAACCGAAGGTGTGCTAAATACCGCATATCTTACGGAAGAAAGCAGGGGATCGCAA	180
OY	181	GACCTTGGCGCTTTTGGAGCGCGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGCTTAC	240
Db	181	GACCTTGGCGCTTTTGGAGCGCGCCGATGTCTGATTAGCTAGTTGGTGAAGTTAAAGCTTAC	240
OY	241	CAGGCGCAGATCATGATGTTGGTCTGAGAGAGACACACGCCACATCGGGACCTGAGACAG	300
Db	241	CAGGCGCTCGAATCAGTATGCTGTGCTGAGAGAGACACACGCCACATCGGGACCTGAGACAG	300
OY	301	GCCCAAGACTCTTACGGGAGAGCAGCAGTGGGGAATTTTGGACATGGGCGCAAGCTGATC	360
Db	301	GCCCAAGACTCTTACGGGAGAGCAGCAGTGGGGAATTTTGGACATGGGGGGAATCCGTGATC	360
OY	361	CAGCAATGCGCGTGAAGTGAAGAAAGGCTTTGGGTTGTAAAGCTCTTTCACTCGAGAGA	420
Db	361	CAGCCATGCGCGGTGAAGTGAAGAAAGGCTTTGGGTTGTAAAGCTCTTTCAAGCCGGAACA	420
OY	421	AAAGGTATCGGTAAATTAATCTGTGACTCATGACGGTATTCACAGAAAGAACCGGCTAAC	480
Db	421	AAACGTCAAGGTAAATTAATCTGTGACTCATGACGGTATTCAGGAAGAAAGAACCGGCTAAC	480
OY	481	TACGTGGCCAGACGCGCGGTAAATACGTAGGATGTCAGACGTTAATCGGAATTAATCGGCGT	540
Db	481	TACGTGCCACGACGCGCGGTAAATACGTAGGATGTCAGACGTTAATCGGAATTAATCGGCGT	540
OY	541	AAAGGTGCGCAGGCGGCTTTGTAACTAGATGTGAATCCCGGGCTTAACTTGGGAAT	600
Db	541	AAAGGTGCGCAGGCGGCTTTGTAACTAGATGTGAATCCCGGGCTTAACTTGGGAAT	600
OY	601	TGCGTTTGAATCTACAAGGCTAGATGTGTGGCAGAGGGAAGTGGAAATTCATGTGTAGCAG	660
Db	601	TGCGTTTGAATCTACAAGGCTAGATGTGTGGCAGAGGGAAGTGGAAATTCACGTGTAGCAG	660
OY	661	TGAAATGCGTAGAGATATGGAAGAAATCATGCATGGCGCAAGCCTCTTGGTAAACAT	720
Db	661	TGAAATGCGTAGAGATGTGGAAGAAATCATGCATGGCGCAAGCCTCCCTTGGTAAACAT	720
OY	721	GAGGCTCATGCAAGAAAGCGTGGGGAGCAAAACAGATTAGATACCTGTATGTCCAGGCC	780
Db	721	GAGGCTCAGGCAAGAAAGCGTGGGGAGCAAAACAGATTAGATACCTGTATGTCCAGGCC	780
OY	781	CTPAAAGATCTGAATATGTTGTGGGCTTATTAAGCTTGGTAAACGAAGCTAACCGGTGA	840
Db	781	CTPAACTATGTAACATATGTTGTGGGCTTAAACGGGCTTGTAAACGACCTTAACGGGTGA	840
OY	841	AGTTGACCGGCTTGGGGATGACGGTGCAGAAATTAAATCTCAAAGAAATTGACGGGACCC	900

[illegible]

```

/organism="Nitrosospira sp. KA98"
/mol_type="genomic DNA"
/db_xref="taxon:289541"
/clone="KA98"
/note="similar to Nitrosospira (Nitrosolobus) multiformis
<1. >1458
/product="16S ribosomal RNA"
rRNA

```

Query Match	88.4%	Score 1287.6	DB 1	Length 1458
Best Local Similarity	93.6%	Pred. No. 3.6e-13		
Matches 186	Conservative	0	Mismatches 89	Indels 5
				Gaps 2

QY	ATTGAAAGCTGGCCGATGCTTTTAAACAATGCAAGTCGAAACGGACGACAGATGCTTCAT	60
Db	1 ATTGAAAGCTGGCCGATGCTTTTAAACAATGCAAGTCGAAACGGACGACAGATGCTTCAT	58
QY	CTGGTGGCGATGGCCGACCGGGTGAATATGATTCGAAACGATTCAGAAAGGGGGGTA	120
Db	59 CTGGTGGCGATGGCCGACCGGGTGAATATGATTCGAAACGATTCCTTAAATGGGGGGTA	118
QY	ACCGATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGGAGGGGATTCGAA	180
Db	119 ACGCACCGAAAGTGTGCTAATATCCGCAATATCTTCGAGAGAAAGAGGGGATTCGAA	178
QY	GACCTTGGCTTTTGGAGCCGCCATGTCTGATTAGTATGTTGGTGAAGGCTTAC	240
Db	179 GACCTTGGCTTTTGGAGCCGCCATGTCTGATTAGTATGTTGGTGAAGGCTTAC	238
QY	CAAGGCCACGATCAAGTATGTGTCTGAGAGGACACACGACCATCTGGGACTGAGACACG	300
Db	239 CAAGCCGTCGATCAAGTATGTGTCTGAGAGGACACGACCATCTGGGACTGAGACACG	298
QY	GCCGAGATCTCTTAACGGAGGACGACGTGGGGAAATTTTGAACATATGGGCGCAAGCTGATC	360
Db	301 GCCGAGATCTCTTAACGGAGGACGACGTGGGGAAATTTTGAACATATGGGCGCAAGCTGATC	358
QY	GCCGAGATCTCTTAACGGAGGACGACGTGGGGAAATTTTGAACATATGGGCGCAAGCTGATC	420
Db	361 GCCGAGATCTCTTAACGGAGGACGACGTGGGGAAATTTTGAACATATGGGCGCAAGCTGATC	418
QY	CAGCATGCGCGGTGATGGAAGAGGCTTGGGTTGTTAAAGCTTTCACTCGAAGAGA	480
Db	421 CAGCATGCGCGGTGATGGAAGAGGCTTGGGTTGTTAAAGCTTTCACTCGAAGAGA	478
QY	AAAGGTTACGGTTAAATTAATCGTACTATGACCGGATGACAGAGAAAGAACCCGGCTTAC	540
Db	419 AAAGGTTACGGTTAAATTAATCGTACTATGACCGGATGACAGAGAAAGAACCCGGCTTAC	538
QY	TACGTGCAGACGCGCGCGTAAATACGTAGGGTGCAGAGCGTTAATTCGAAATTAACCTGGCGGT	600
Db	481 TACGTGCAGACGCGCGCGTAAATACGTAGGGTGCAGAGCGTTAATTCGAAATTAACCTGGCGGT	598
QY	AAAGGTTACGGTTAAATTAATCGTACTATGACCGGATGACAGAGAAAGAACCCGGCTTAC	660
Db	541 AAAGGTTACGGTTAAATTAATCGTACTATGACCGGATGACAGAGAAAGAACCCGGCTTAC	658
QY	TGCGTTTGAATCTCAAGGCTAGAGTGTGGACAGAGGAGTGTGAATTCATAGTGTAGCAG	720
Db	601 TGCGTTTGAATCTCAAGGCTAGAGTGTGGACAGAGGAGTGTGAATTCATAGTGTAGCAG	718
QY	TGAATATGCTGATGAGATGTGAGAGAACACCGATGGCGAAGGGCAGCCCTCTGGGTTAACCT	780
Db	659 TGAATATGCTGATGAGATGTGAGAGAACACCGATGGCGAAGGGCAGCCCTCTGGGTTAACCT	778
QY	GAGCCTACGACGAAAGCGTGGGGAGCAAAACAGATTAAGTACCCTGTAGTCCAGCACC	840
Db	721 GAGCCTACGACGAAAGCGTGGGGAGCAAAACAGATTAAGTACCCTGTAGTCCAGCACC	838
QY	CTAAACGATGTCAACTAGTTGTGGGCTTTAAGGCTTGGTTAAGCAAGCTTAACGCGTGA	900
Db	779 CTAAACGATGTCAACTAGTTGTGGGCTTTAAGGCTTGGTTAAGCAAGCTTAACGCGTGA	898
QY	AGTTGACCGCTGGGGAGTACGCTGTGCAAGATTTAAATCTCAAAAGGAATTTGACGGGAGACC	960
Db	841 AGTTGACCGCTGGGGAGTACGCTGTGCAAGATTTAAATCTCAAAAGGAATTTGACGGGAGACC	958
QY	AGTTGACCGCTGGGGAGTACGCTGTGCAAGATTTAAATCTCAAAAGGAATTTGACGGGAGACC	1020
Db	839 AGTTGACCGCTGGGGAGTACGCTGTGCAAGATTTAAATCTCAAAAGGAATTTGACGGGAGACC	1018

QY	901	GCACAAGCGGGTGAAATTATGTGATTTAAATTCATGCAACGCGAAAAAAGCTTACCTACCCCTT	960
Db	899	GCAACAAGCGGGTGAAATTATGTGATTTAAATTCATGCAACGCGAAAAAAGCTTACCTACCCCTT	958
QY	961	GACATGTAGCGAATTTTCTAGAGATAGATTATGTC---TTGCGGAAAGCGTAAACAGGTG	1012
Db	959	GACATGTATCCGAAACCCGCGCGAGAGGTGGGTGTCCCGAAAGGGAACGGCAACAGGTG	1018
QY	1018	CTGCATGGCTGTGCTCAAGCTGTGTGCTGTGAGATGTGTGGTTAAATTCCTCCGCAACGAGCGCA	1077
Db	1019	CTGCATGGCTGTGCTCAAGCTGTGTGCTGTGAGATGTGTGGTTAAATTCCTCCGCAACGAGCGCA	1078
QY	1078	ACCCGTGTCATTAAATTTGGCATCATTTGGTTGGGCACTTAAATGAGACTGCCGGTACAA	1133
Db	1079	ACCCGTGTCATTAAATTTGGCATCATTTGAGTGGGCACTTAAATGAAACTGCCGGTACAA	1133
QY	1138	CCGAGAGAGGTGGGAGATGACGTCAAGTCTCATGCGCCCTTAATGGGTAGGGCTTCACAG	1197
Db	1139	CCGAGAGAGGTGGGAGATGACGTCAAGTCTCATGCGCCCTTAATGGGTAGGGCTTCACAG	1198
QY	1198	TAAATCAATGGCGCGTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTATCTCAGAAAGC	1257
Db	1199	TAAATCAATGGCGCGTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTATCTCAGAAAGC	1258
QY	1258	GCGTCGTATGTCGGATGGGAGTCTGCAACTGACCTCGTGAAAGTCGGAATCGCTAGTAT	1312
Db	1259	GCGTCGTATGTCGGATGGGAGTCTGCAACTGACCTCGTGAAAGTCGGAATCGCTAGTAT	1318
QY	1318	CGCGGATCAGCATGTCGCGGTGAATAGATTCCCGGACTTTGTACACAACCGCGCGTCAAC	1377
Db	1319	CGCGGATCAGCATGTCGCGGTGAATAGATTCCCGGACTTTGTACACAACCGCGCGTCAAC	1378
QY	1378	CATGGAGATGGGTTTCAACCAAGACAGGTAGTCTTAACCGTAAAGAGGGCGCTTGCACAG	1437
Db	1379	CATGGAGATGGGTTTCAACCAAGACAGGTAGTCTTAACCGCAAGAGAGGGCGCTTGCACAG	1438
QY	1438	TGAGATTCAATCTGGGGTG 1457	
Db	1439	TGAGATTCAATCTGGGGTG 1458	

RESULT	59
LOCUS	CQ796910
DEFINITION	Sequence 3 from Patent WO2004026772.
ACCESSION	CQ796910
VERSION	CQ796910.1
KEYWORDS	GI:46408536
SOURCE	.
ORGANISM	unidentified unidentified unclassified sequences.
REFERENCE	1
AUTHORS	Hovaneec,T.A.
TITLE	Alpha-oxidizing bacteria and methods of using and detecting thesame
JOURNAL	Patent: WO 2004026772-A 3 01-APR-2004;
FEATURES	Aquaria Inc. (US)
source	Location/Qualifiers 1..1458 /organism="unidentified" /mol_type="unassigned DNA" /db_xref="taxon:32644" /note="AOB Type B rDNA clones 16S rDNA"
ORIGIN	

	Query Match	88.4%	Score 1287.6;	DB 6;	Length 1458;
	Best Local Similarity	93.6%;	Pred. No. 3.6e-13;		
	Matches 1366;	Conservative 0;	Mismatches 89;	Indels 5;	Gaps 2;
OY	1 ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCCAAACGCCAGCAGCGAATGTCTTGAT	60			
Db	1 ATTTAACGCTGGCGGCATGCTTTACACATGCAAGTCCAAACGCCAGCAGCGAATGTCTTGAT--	58			

QY 61 CTGTGGCGAGTGGCGGAGCGGTGAGTAAATGATCGGAAAGTATCCAGAAAGAGGGGGTAA 120  
 DB 59 CTGTGGCGAGTGGCGGAGCGGTGAGTAAATGATCGGAAAGTATCCAGAAAGAGGGGGTAA 118  
 QY 121 AGCGATCGAAAGATGTCTTAATCCGCATTAATCTTAAGAGAGAAAGCGGGGATGAGAA 180  
 DB 119 AGCGATCGAAAGATGTCTTAATCCGCATTAATCTTAAGAGAGAAAGCGGGGATGAGAA 178  
 QY 181 GACCTTGCGCTTTGGAGGGCGCGATGTCTGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 240  
 DB 179 GACCTTGCGCTTTGGAGGGCGCGATGTCTGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 238  
 QY 241 CAAGCGAGCATGATGATTGGTCTGAGAGAGACGACGACACTGGGACTGAGACAGC 300  
 DB 239 CAAGCGAGCATGATGATTGGTCTGAGAGAGACGACGACACTGGGACTGAGACAGC 298  
 QY 301 GCCCAAGCTCTTACGGGAGGACAGATGGGGAAATTTTGGACATATGGGCGGCAAGCTGATC 360  
 DB 299 GCCCAAGCTCTTACGGGAGGACAGATGGGGAAATTTTGGACATATGGGCGGCAAGCTGATC 358  
 QY 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTCTTTCAGTCAGAGAGA 420  
 DB 359 CAGCAATGCCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTCTTTCAGTCAGAGAGA 418  
 QY 421 AAAGTTACCGTAAATTAATGCTGATCATGACGGTATCGACAGAAAGACACCGGCTTAC 480  
 DB 419 AAAGTTACCGTAAATTAATGCTGATCATGACGGTATCGACAGAAAGACACCGGCTTAC 478  
 QY 481 TACGTGCACAGACCGCGGTAAATACGTAGGCTGCAAGCTTAAATTCGAATTCAGGCGCT 540  
 DB 479 TACGTGCACAGACCGCGGTAAATACGTAGGCTGCAAGCTTAAATTCGAATTCAGGCGCT 538  
 QY 541 AAAGGGTGGCGGCGCTTGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
 DB 539 AAAGGGTGGCGGCGCTTGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 598  
 QY 601 TGGCTTTGAAACTACAGAGCTAGAGTGTGCGAGAGGAGGTGGAATTCATGTGTACAG 660  
 DB 599 TGGCTTTGAAACTACAGAGCTAGAGTGTGCGAGAGGAGGTGGAATTCATGTGTACAG 658  
 QY 661 TGAATTCGTAGATATGGAAGACATCGATGGCGAAGGCGCTTCTGGGTTAAACT 720  
 DB 659 TGAATTCGTAGATATGGAAGACATCGATGGCGAAGGCGCTTCTGGGTTAAACT 718  
 QY 721 GAGCTCATGACGAAAGCGTGGGAGCAACAGATTAATGATTAATGATTAATGATTAATGAT 780  
 DB 719 GAGCTCATGACGAAAGCGTGGGAGCAACAGATTAATGATTAATGATTAATGATTAATGAT 778  
 QY 781 CTAAACGATGCACTAGTTGTTGGGCTTATTAAGCTTGGTAAAGAGCTTAACGCGTGA 840  
 DB 779 CTAAACGATGCACTAGTTGTTGGGCTTATTAAGCTTGGTAAAGAGCTTGAACGCGTGA 838  
 QY 841 AGTTGACCGCTGGGAGATGAGTGTGCGAAGATTAATCAAGAGATTAAGCGGGAGCC 900  
 DB 839 AGTTGACCGCTGGGAGATGAGTGTGCGAAGATTAATCAAGAGATTAAGCGGGAGCC 898  
 QY 901 GCAACAGCGGTGATTAATGAGTAAATGATGCAACGCGAATACTTAATCTTACCCTT 960  
 DB 899 GCAACAGCGGTGATTAATGAGTAAATGATGCAACGCGAATACTTAATCTTACCCTT 958  
 QY 961 GACATGATGCAATTTTCTAGAGATGATTAATGATGCAATTTTCTAGAGATGATTAATGAT 1017  
 DB 959 GACATGATGCAATTTTCTAGAGATGATTAATGATGCAATTTTCTAGAGATGATTAATGAT 1018  
 QY 1018 CTGATGAGCTGTCTGATGCTGTGTCTGAGATGTTGGGTTAAAGTCCCGCAAGAGCGCA 1077  
 DB 1019 CTGATGAGCTGTCTGATGCTGTGTCTGAGATGTTGGGTTAAAGTCCCGCAAGAGCGCA 1078  
 QY 1078 ACCCTGTCATTAATGTCATCATTTGGTGGGCACTTAATGAGACTGCGGTGACAA 1137  
 DB 1079 ACCCTGTCATTAATGTCATCATTTGGTGGGCACTTAATGAGACTGCGGTGACAA 1138  
 QY 1138 CCGAGGAAGGTGGGATGAGCTCAAGTCTCATGGCCCTTAATGGGTTGAGGCTTCAACG 1197

DB 1139 CCGAGGAAGGTGGGATGACCTCAAGTCTCTATGCGCTTAATGGGTGAGGCTTCAACG 1198  
 QY 1198 TAATCAATGAGCGCTACAGAGGTTGCCAACCCCGAGGGGAGACTAATCTCAAGAAC 1257  
 DB 1199 TAATCAATGAGCGCTACAGAGGTTGCCAACCCCGAGGGGAGACTAATCTCAAGAAC 1258  
 QY 1258 GCGTGTAGTCCGATCGGATGCTGCACTGCACTCCGGAATGGGAATGGCTAGTAAT 1317  
 DB 1259 GCGTGTAGTCCGATCGGATGCTGCACTGCACTCCGGAATGGGAATGGCTAGTAAT 1318  
 QY 1318 CCGGATCAGCATGTGCGGTAATACGTTCCGGGCTTGTACACCCCGCTGACAC 1377  
 DB 1319 CCGGATCAGCATGTGCGGTAATACGTTCCGGGCTTGTACACCCCGCTGACAC 1378  
 QY 1378 CATGGAGTGGGTTTCCAGAGACAGTATCTTAACGTTAAAGAGGGGCTTTGCCACG 1437  
 DB 1379 CATGGAGTGGGTTTCCAGAGACAGTATCTTAACGTTAAAGAGGGGCTTTGCCACG 1438  
 QY 1438 TGAATTCATGACTGGGGTG 1457  
 DB 1439 TGAATTCATGACTGGGGTG 1458  
 RESULT 60  
 CS089156 1458 bp DNA linear PAT 25-MAY-2005  
 LOCUS CS089156  
 DEFINITION Sequence 3 from Patent EP1502948.  
 ACCESSION CS089156  
 VERSION CS089156.1 GI:66714440  
 KEYWORDS  
 SOURCE unidentifed  
 ORGANISM unidentifed  
 unclassified.  
 REFERENCE 1  
 AUTHORS Hovanec, T.A. and Burrell, P.C.  
 TITLE Ammonia-oxidizing bacteria  
 JOURNAL Patent: EP 1502948-A 3 02-FEB-2005;  
 Aquaria Inc. (US)  
 FEATURES  
 source 1..1458  
 /organism="unidentifed"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3264"  
 /note="Description of Unknown Organism: AOB Type B  
 ammonia-oxidizing bacterium represented by R3 clones"  
 ORIGIN  
 Query Match 88.4%; Score 1287.6; DB 6; Length 1458;  
 Best Local Similarity 93.6%; Pred. No. 3.6e-13;  
 Matches 1366; Conservative 0; Mismatches 89; Indels 5; Gaps 2;  
 QY 1 ATTGAACGCTGGCGGCGATGCTTATACATCAATGCGAAGCGGAGCATGTCAT 60  
 DB 1 ATTGAACGCTGGCGGCGATGCTTATACATCAATGCGAAGCGGAGCATGTCAT 58  
 QY 61 CTGTGGCGAGTGGCGGAGCGGTGAGTAAATGATCGGAACGTAATCCAGAAAGAGGGGCTA 120  
 DB 59 CTGTGGCGAGTGGCGGAGCGGTGAGTAAATGATCGGAACGTAATCCAGAAAGAGGGGCTA 118  
 QY 121 AGCGATCGAAAGATGTCTTAATCCGCATTAATCTTAAGAGAGAAAGCGGGGATGAGAA 180  
 DB 119 AGCGATCGAAAGATGTCTTAATCCGCATTAATCTTAAGAGAGAAAGCGGGGATGAGAA 178  
 QY 181 GACCTTGCGCTTTGGAGGGCGCGATGTCTGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 240  
 DB 179 GACCTTGCGCTTTGGAGGGCGCGATGTCTGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 238  
 QY 241 CAAGCGAGCATGATGATTGGTCTGAGAGAGACGACGACACTGGGACTGAGACAGC 300  
 DB 239 CAAGCGAGCATGATGATTGGTCTGAGAGAGACGACGACACTGGGACTGAGACAGC 298  
 QY 301 GCCCAAGCTCTTACGGGAGGACAGATGGGGAAATTTTGGACATATGGGCGGCAAGCTGATC 360

Db 299 GCCCAACTCTTACGGGAGGACAGTGGGAAATTTTGACATGGGGAAACCTGATC 358  
Qy 361 CAGCAATGCCGCGTGAAGAGAGCCCTTCGGGTTGTAAGCTCTTTCAGTCGAGAGA 420  
Db 359 CAGCCATGCCGCGTGAAGAGAGCCCTTCGGGTTGTAAGCTCTTTCAGTCGAGAGA 418  
Qy 421 AAAGTTAGGTTAAATCTGATCATGACGGTTCGACAGAGAGACACCGGCTAAC 480  
Db 419 AACGGTACGGCTTAATACCGTACTACTGACGGTACCGAGAGAGAGACCGGCTAAC 478  
Qy 481 TACGTCCACAGACCCCGGTAAATACGAGGTCGAAGCGTTAATCGGAATTACTGGGCGT 540  
Db 479 TACGTCCACAGACCCCGGTAAATACGAGGTCGAAGCGTTAATCGGAATTACTGGGCGT 538  
Qy 541 AAAGGTGCGCAGGCGGCTTTTGAAGTACAGATGTAATCCCGGGCTTAACTGGGAAT 600  
Db 539 AAAGCGTGGCAGGCGGCTTTTGAAGTACAGATGTAATCCCGGGCTTAACTGGGAAT 598  
Qy 601 TCGGTTGAAACTTACAAAGCTAGAGTGGCAGAGGAGGTGGAATTCATGTTAGCAG 660  
Db 599 TCGGTTGAAACTTACAAAGCTAGAGTGGCAGAGGAGGTGGAATTCATGTTAGCAG 658  
Qy 661 TGAATGCGTAGAGATATGAAAGAAATCATGATGCGAAGGCGCTCTGGGTTAACT 720  
Db 659 TGAATGCGTAGAGATATGAAAGAAATCATGATGCGAAGGCGCTCTGGGTTAACT 718  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTGTTAGTCCAGCC 780  
Db 719 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTGTTAGTCCAGCC 778  
Qy 781 CTAAACGATGTAACATGTTGTTGGGCTTATTAAGCTTGTGTAACGAACTAACCGGTGA 840  
Db 779 CTAAACGATGTAACATGTTGTTGGGCTTATTAAGCTTGTGTAACGAACTAACCGGTGA 838  
Qy 841 AGTTGACCGCTGGGAGTACCGTCCGAAAGTTAACTCAAGAAATGACGGGAGCC 900  
Db 839 AGTTGACCGCTGGGAGTACCGTCCGAAAGTTAACTCAAGAAATGACGGGAGCC 898  
Qy 901 GCACAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAACCTTACCTACCTT 960  
Db 899 GCACAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAACCTTACCTACCTT 958  
Qy 961 GACATGACGAAATTTTGAAGATAGATTAGTGC---TTGGGAAACGTTAACACAGGTG 1017  
Db 959 GACATGACGAAACCCCGGAGAGTGGGTGCCCCGAAAGGAGCGGTAAACACAGGTG 1018  
Qy 1018 CTGCATGCGCTGCTGACGCTCGTGTGAGATGTTGGGTTAAAGTCCGCAACGAGCGCA 1077  
Db 1019 CTGCATGCGCTGCTGACGCTCGTGTGAGATGTTGGGTTAAAGTCCGCAACGAGCGCA 1078  
Qy 1078 ACCCTTGTCAATTAATGCAATCAATTTGGTGGGCACTTTAATGAGACTGCGGTGCAAA 1137  
Db 1079 ACCCTTGTCAATTAATGCAATCAATTTGGTGGGCACTTTAATGAGACTGCGGTGCAAA 1138  
Qy 1138 CCGGAGGAAGGTGGGAGTACGTCAGTCTCAATGCGCTTAAAGGTAGGGCTTCAACAG 1197  
Db 1139 CCGGAGGAAGGTGGGAGTACGTCAGTCTCAATGCGCTTAAAGGTAGGGCTTCAACAG 1198  
Qy 1198 TAAATCAATGCGCGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1257  
Db 1199 TAAATCAATGCGCGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1258  
Qy 1258 GCGTGTAGTCCGGAATCGAGTCTGCAACTCGACTCGTAAGTGGAAATCGCTAGTAAT 1317  
Db 1259 GCGTGTAGTCCGGAATCGAGTCTGCAACTCGACTCGTAAGTGGAAATCGCTAGTAAT 1318  
Qy 1318 CCGGATCAGCATGTCGCGGTAAATACGTTCCCGGTCTTGTACACACCGCCCGTACAC 1377  
Db 1319 CCGGATCAGCATGTCGCGGTAAATACGTTCCCGGTCTTGTACACACCGCCCGTACAC 1378  
Qy 1378 CATGGAGTGGGTTTACACAGAGCAGTATGTTAAACCGTAAAGAGGGCGCTTGCCACGG 1437

Db 1379 CATGGAGTGGGTTTACACAGAGCAGTATGTTAAACCGAAGAGAGGGCGCTTGCCACGG 1438  
Qy 1438 TGAATTCATGACTGGGTTG 1457  
Db 1439 TGAATTCATGACTGGGTTG 1458  
RESULT 61  
AX316094  
LOCUS AX316094 1458 bp DNA linear PAT 14-DEC-2001  
DEFINITION Sequence 3 from Patent WO0190312.  
ACCESSION AX316094  
VERSION AX316094.1 GI:17899285  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1  
AUTHORS Hovanec,T.A. and Burrell,P.C.  
TITLE Ammonia-oxidizing bacteria  
JOURNAL Patent: WO 0190312-A 3 29-NOV-2001;  
AQUARIA, INC. (US)  
FEATURES  
source  
1..1458  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="AOB type B ammonia-oxidizing bacterium represented  
by R3 clones"  
ORIGIN  
Query Match 88.4%; Score 1287.6; DB 6; Length 1458;  
Best Local Similarity 93.6%; Pred. No. 3.6e-13;  
Matches 136; Conservative 0; Mismatches 89; Indels 5; Gaps 2;  
Qy 1 ATTTGAACGCTGGCGGATGCTTTACATGCAATGCAAGTCGACGAGCGATGCTTGAT 60  
Db 1 ATTTGAACGCTGGCGGATGCTTTACATGCAATGCAAGTCGACGAGCGGAGCAAC-- 58  
Qy 61 CTGTGGCGAGTGGCGGACGGGTGATGATGATGCAAGCTATTCAGAAAGGGGGGTA 120  
Db 59 CTGTGGCGAGTGGCGGACGGGTGATGATGATGCAAGCTATTCAGAAAGGGGGGTA 118  
Qy 121 AGCGATGAAAGTGTGTAATACCGCATATATCTTAAGAGAGAGAAAGCAGGGGATCGAAA 180  
Db 119 AGCGATGAAAGTGTGTAATACCGCATATATCTTAAGAGAGAGAAAGCAGGGGATCGAAA 178  
Qy 179 GACCTTGGGCTTGGAGCGGCGCATGTCGATTAGCTAGTTGGTGAAGTAAAGCTTAC 238  
Qy 241 CAAAGCGCATCAGTACGATGTTGCTGAGAGAGACGACGACACTGCGACTGAGACAG 300  
Db 239 CAAAGCGCATCAGTACGATGTTGCTGAGAGAGACGACGACACTGCGACTGAGACAG 298  
Qy 301 GCCCAGACTCCTTACGGGAGGAGCAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360  
Db 299 GCCCAGACTCCTTACGGGAGGAGCAGTGGGAAATTTTGGACAAATGGGCGAAACCTGATC 358  
Qy 361 CAGCAATGCCGCGTGAAGAGAGCCCTTCGGGTTGTAAGCTCTTTCAGTCGAGAGA 420  
Db 359 CAGCAATGCCGCGTGAAGAGAGCCCTTCGGGTTGTAAGCTCTTTCAGTCGAGAGA 418  
Qy 421 AAAGTTAGGTTAAATCTGATCATGACGGTTCGACAGAGAGACACCGGCTAAC 480  
Db 419 AAAGTTAGGTTAAATCTGATCATGACGGTTCGACAGAGAGACACCGGCTAAC 478  
Qy 481 TACGTCCACAGACCCCGGTAAATACGAGGTCGAAGCGTTAATCGGAATTACTGGGCGT 540  
Db 479 TACGTCCACAGACCCCGGTAAATACGAGGTCGAAGCGTTAATCGGAATTACTGGGCGT 538  
Qy 541 AAAGGTGCGCAGGCGGCTTTGTAAGTACAGATGTAATCCCGGGCTTAACTGGGAAT 600

```

Db      539 AAAGCGTGGCAGCGGTTTTTTGTAAGTCAGATGTAAAGCCCCGGGCTTAACCTGGGAAC 598
Qy      601 TGGGTTGAAACTACAAGGCTAGAGTGTGGCAGAGGAGGTGGAAATTCATGTGTAGCAG 660
Db      599 TGGGTTGAAACTACAAGGCTAGAGTGTGGCAGAGGAGGTGGAAATTCACGTGTAGCAG 658
Qy      661 TGAATATGCTAGAGATATGAGAAACATCGATGGCGAAGCGACCTCTGGGTTAACT 720
Db      659 TGAATATGCTAGAGATATGAGAAACATCGATGGCGAAGCGACCTCTGGGTTAACT 718
Qy      721 GACGCTCATGACAGAAAGCTGGGGAGCAAAACGAAATTAATCCTGTAGTCCAGCC 780
Db      719 GACGCTCATGACAGAAAGCTGGGGAGCAAAACGAAATTAATCCTGTAGTCCAGCC 778
Qy      781 CTAAACGATGTCAACTAGTTGTTGGGCTTAATTAAGGCTTGTATACGAAGCTTAACGCGTGA 840
Db      779 CTAAACGATGTCAACTAGTTGTTGGGCTTAATTAAGGCTTGTATACGAAGCTTAACGCGTGA 838
Qy      841 AGTTGACCGGCTGGGGAGTACGGTCCGCAAGATTAACTCAAGAAATTGACGGGGAACC 900
Db      839 AGTTGACCGGCTGGGGAGTACGGTCCGCAAGATTAACTCAAGAAATTGACGGGGAACC 898
Qy      901 GCAACAGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAACCTTAACCTACCTT 960
Db      899 GCAACAGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAACCTTAACCTACCTT 958
Qy      961 GAAATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTGGGAAACGCTTACACAGGTG 1017
Db      959 GAAATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTGGGAAACGCTTACACAGGTG 1018
Qy      1018 CTGCATAGCTGTGCTGACGTCTGTCTGTGATGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1077
Db      1019 CTGCATAGCTGTGCTGACGTCTGTCTGTGATGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1078
Qy      1078 ACCCTGTCTAATTAATTCGATCAATTTGGTGGGCACTTTAATGAGACTGCGCGGTGACAA 1137
Db      1079 ACCCTGTCTAATTAATTCGATCAATTTGGTGGGCACTTTAATGAGACTGCGCGGTGACAA 1138
Qy      1138 CCGGAGGAAGGTGGGGAGTGAAGTCGATGCTCATGAGCCCTTAATGAGGAGGCTTACACAG 1197
Db      1139 CCGGAGGAAGGTGGGGAGTGAAGTCGATGCTCATGAGCCCTTAATGAGGAGGCTTACACAG 1198
Qy      1198 TAATTAACATGCGCGGTACAGAGGGTTGCAACCCGCGAGGGGAGCTAATCTCGAAGAAC 1257
Db      1199 TAATTAACATGCGCGGTACAGAGGGTTGCAACCCGCGAGGGGAGCTAATCTCGAAGAAC 1258
Qy      1258 GCGTCTGATGTCGGGATGGAAGTCTGCACTCGTGAAGTTCGGAATTCGCTAGTAAT 1317
Db      1259 GCGTCTGATGTCGGGATGGAAGTCTGCACTCGTGAAGTTCGGAATTCGCTAGTAAT 1318
Qy      1318 CCGGAGTACACATGTCGGGAGTAAGTTCGCGGGCTTGTATACACCGCGCTGACAC 1377
Db      1319 CCGGAGTACACATGTCGGGAGTAAGTTCGCGGGCTTGTATACACCGCGCTGACAC 1378
Qy      1378 CATGGAGTGGGTTTACACAGAACAGAGTCTTAACCGCAAGAGGCGCTTTCACAGG 1437
Db      1379 CATGGAGTGGGTTTACACAGAACAGAGTCTTAACCGCAAGAGGCGCTTTCACAGG 1438
Qy      1438 TGAGATTCAATGACTGGGGTG 1457
Db      1439 TGAGATTCAATGACTGGGGTG 1458

```

```

RESULT 62
LOCUS      AY123796              1497 bp    DNA          linear    BCT 23-SEP-2003
DEFINITION Nitrosospiira sp. L115 16S ribosomal RNA gene, partial sequence.
ACCESSION  AY123796
VERSION     AY123796.1 GI:24474425
KEYWORDS
SOURCE      Nitrosospiira sp. L115
ORGANISM    Nitrosospiira sp. L115
            Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;

```

```

REFERENCE 1 (bases 1 to 1497)
AUTHORS   Purthold,U., Wagner,M., Timmermann,G., Pommerening-Roser,A. and
           Koops,H.-P.
TITLE      16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial
           ammonia-oxidizing isolates: extension of the dataset and proposal
           of a new lineage within the nitrosomonads
JOURNAL    Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)
PUBMED     13130037
REFERENCE 2 (bases 1 to 1497)
AUTHORS   Purthold,U., Timmermann,G., Koops,H.-P., Pommerening-Roser,A. and
           Wagner,M.
TITLE      Direct Submission
JOURNAL    Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische
           Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES   Location/Qualifiers
            source          1..1497
                           /organism="Nitrosospiira sp. L115"
                           /mol_type="genomic DNA"
                           /strain="L115"
                           /db_xref="taxon:174928"
                           <1..>1497
                           /product="16S ribosomal RNA"
ORIGIN
Query Match      88.4%; Score 1287.6; DB 1; Length 1497;
Best Local Similarity 93.6%; Pred. No. 3,5e-13;
Matches 1366; Conservative 0; Mismatches 89; Indels 5; Gaps 2;

```

```

Qy      1 ATTGAACGTGGCGGAGTCTTTTACACATGCAAGTCAACGCGACGACGAGATCTTGAT 60
Db      3 ATTGAACGTGGCGGAGTCTTTTACACATGCAAGTCAACGCGACGACGAGATCTTGAT 60
Qy      61 CTGGTGGCAGATGGCGAGCGGAGTGAATGCAATCGGAACGTAATCCAGAAAGAGGGGAGTA 120
Db      61 CTGGTGGCAGATGGCGAGCGGAGTGAATGCAATCGGAACGTAATCCAGAAAGAGGGGAGTA 120
Qy      121 ACGCATCAAGAAAGTGTCTAATACCGCATTAATCTTAAGAGAGAAACGAGGAGTGCAG 180
Db      121 ACGCATCAAGAAAGTGTCTAATACCGCATTAATCTTAAGAGAGAAACGAGGAGTGCAG 180
Qy      181 GACCTTGGCGCTTTTGGAGGGCGCGATGTCGATTAGCTAGTGGTGGGTTAAAGGCTTAC 240
Db      181 GACCTTGGCGCTTTTGGAGGGCGCGATGTCGATTAGCTAGTGGTGGGTTAAAGGCTTAC 240
Qy      241 CAAGCGACATGATGATGTTGTTCTGAGAGAGACCAAGCACTGGGACTGAGACAG 300
Db      241 CAAGCGACATGATGATGTTGTTCTGAGAGAGACCAAGCACTGGGACTGAGACAG 300
Qy      301 GCCCAGACTTCCTAAGCGGAGGACGACATGGGGAATTTTGGCAATGGGCGCAACCTGATC 360
Db      301 GCCCAGACTTCCTAAGCGGAGGACGACATGGGGAATTTTGGCAATGGGCGCAACCTGATC 360
Qy      361 CAGCAATGCCGCTGATGAGTAAGAAAGCTTCGGGTTGTAAGCTCTTTCAGTGAGAGA 420
Db      361 CAGCAATGCCGCTGATGAGTAAGAAAGCTTCGGGTTGTAAGCTCTTTCAGTGAGAGA 420
Qy      421 AAAGGTTACGGTAAATTAATCTGACTCATGACGGTATTCACAGAAAGAACCGGCTAAC 480
Db      421 AAAGGTTACGGTAAATTAATCTGACTCATGACGGTATTCACAGAAAGAACCGGCTAAC 480
Qy      481 TAGGTGCAACAGCCGCGGTAATACGATGGGTCAAGCGTTAATCGGAATTAACCTGGCGGT 540
Db      481 TAGGTGCAACAGCCGCGGTAATACGATGGGTCAAGCGTTAATCGGAATTAACCTGGCGGT 540
Qy      541 AAAGGTCGCGACGCGGCTTTTGAAGTCAATGTAATCCCGGCTTAACCTGGGAAT 600
Db      541 AAAGGTCGCGACGCGGCTTTTGAAGTCAATGTAATCCCGGCTTAACCTGGGAAT 600
Qy      601 TGGGTTGAAACTACAAGGCTAGAGTGTGCAAGAGGAGGTGAAATTCATGTGTAGCAG 660
Db      601 TGGGTTGAAACTACAAGGCTAGAGTGTGCAAGAGGAGGTGAAATTCATGTGTAGCAG 660

```

[illegible]

JOURNAL	of a new lineage within the nitrospinales
PUBMED	Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)
REFERENCE	13130037
AUTHORS	2 (bases 1 to 1497)
	Pukthold, U., Timmermann, G., Koop, H.-P., Pommerening-Roeser, A. and Wagner, M.
TITLE	Direct Submission
JOURNAL	Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES	Location/Qualifiers
source	1. 1497
	/organism="Nitrospira sp. N15"
	/mol_type="genomic DNA"
	/strain="N15"
	/db_xref="taxon:200120"
	<1. .>1497
	/product="16S ribosomal RNA"
ORIGIN	
Query Match	88.4%; Score 1287.6; DB 1; Length 1497;
Best Local Similarity	93.6%; Pred. No. 3.5e-13;
Matches 1366; Conservative	0; Mismatches 89; Indels 5; Gaps 2
Db	1 ATGGAAGCTGCGCGCATGCTTTACATCGAAGTCGAACGGACGACGAGTGTGCT 60
Qy	
Db	3 ATTGAAGCTGCGCGCATGCTTTACATCGAAGTCGAACGGACGACGAGTGTGCT 60
Qy	61 CTGCTGGCGATGCGCGCATGCTTTACATCGAAGTCGAACGGACGACGAGTGTGCT 120
Db	61 CTGCTGGCGATGCGCGCATGCTTTACATCGAAGTCGAACGGACGACGAGTGTGCT 120
Qy	121 ACGCATGGAAGATGTGCTTAATACCGCATTAATCTTAAGAGGAAGCAGGGATCGAA 180
Db	121 ACGCATGGAAGATGTGCTTAATACCGCATTAATCTTAAGAGGAAGCAGGGATCGAA 180
Qy	181 GACCTGGCGCTTTGGAGCGCGCGATGCTAGTAGCTAGTGGTGGGTAAAGGCTAC 240
Db	181 GACCTGGCGCTTTGGAGCGCGCGATGCTAGTAGCTAGTGGTGGGTAAAGGCTAC 240
Qy	241 CAAAGCGCATGCTAGTAGTGGTCTGAGAGACGACCGACCATGCGGACTGAGACAG 300
Db	241 CAAAGCGCATGCTAGTAGTGGTCTGAGAGACGACCGACCATGCGGACTGAGACAG 300
Qy	301 GCCCAGACTCTTACCGGAGGACGACGATGGGAAATTTGGACATGGCGCGAAAGCTGATC 360
Db	301 GCCCAGACTCTTACCGGAGGACGACGATGGGAAATTTGGACATGGCGCGAAAGCTGATC 360
Qy	361 CAGCAATGCCCGTGAAGTGAAGAGCCCTTGGGTGTAAGCTCTTCACTCGAGGAAG 420
Db	361 CAGCAATGCCCGTGAAGTGAAGAGCCCTTGGGTGTAAGCTCTTCACTCGAGGAAG 420
Qy	421 AAAGTTACGGTAAATAATCGTCACTGACCGTATGACAGAGAAGACCGGCTAAC 480
Db	421 AAAGTTACGGTAAATAATCGTCACTGACCGTATGACAGAGAAGACCGGCTAAC 480
Qy	481 TAGCTGCCAGACCGCGGTAATACGTAAGGCTGCAAGCTTAATCGAATTACTGGCGT 540
Db	481 TAGCTGCCAGACCGCGGTAATACGTAAGGCTGCAAGCTTAATCGAATTACTGGCGT 540
Qy	541 AAAGGTGCGGACCGCGCTTTGTAAGTCAAGTGAATATCCCGGGCTTAATCTGGGAAT 600
Db	541 AAAGGTGCGGACCGCGCTTTGTAAGTCAAGTGAATATCCCGGGCTTAATCTGGGAAT 600
Qy	601 TGCGTTGAATCTCAAGGCTAGAGTGGGACGAGAGGAGTGAATTCATCTGTAGAG 660
Db	601 TGCGTTGAATCTCAAGGCTAGAGTGGGACGAGAGGAGTGAATTCATCTGTAGAG 660
Qy	661 TGAATCGTGAAGATAGAAACATCGATGGCGAAGCAGCTCTTGAGTTAACT 720
Db	661 TGAATCGTGAAGATAGAAACATCGATGGCGAAGCAGCTCTTGAGTTAACT 720
Qy	721 GAGCTCATGACGAAGCGTGGGAGCAACAGATTAAATCCCTGTAGTCCACCC 780
Db	721 GAGCTCATGACGAAGCGTGGGAGCAACAGATTAAATCCCTGTAGTCCACCC 780

Db 721 GACGCTCAGGACGAAAGCGTGGGAGCAAAAGATTAGATACCTGTGTAATCCACGCC 780  
Qy 781 CTAAGCATGTCACTAGTTGTTGGCCCTTATTAGGCTTGTTACGAAACGCGTGA 840  
Db 781 CTAAGCATGTCACTAGTTGTTGGCTTTAAACGACTTGTAAACGCGTGA 840  
Qy 841 AGTTGACCGCTGGGAGTACGCTGCAAGATTAAACTCAAGATTGACGCGGACCC 900  
Db 841 AGTTGACCGCTGGGAGTACGCTGCAAGATTAAACTCAAGATTGACGCGGACCC 900  
Qy 901 GCAAGACCGGTGATTTATGTGATTTAATTCGATGCAACGCGAAACCTTACCTACCTT 960  
Db 901 GCAAGACCGGTGATTTATGTGATTTAATTCGATGCAACGCGAAACCTTACCTACCTT 960  
Qy 961 GACATGTACGATTTTCTAGAGATAGATTAGTGCT---TTCCGGAACGCTTAAACACAGTG 1017  
Db 961 GACATGTACGAGAACCCCGGAGAGTGGGTGTGCCGAAAGGAGCGGTAAACACAGTG 1020  
Qy 1018 CTGCATGCGTGTGTCAGCTCGTGTGTAAGATTGGTTAAAGTCCCGCAACGAGCGCA 1077  
Db 1021 CTGCATGCGTGTGTCAGCTCGTGTGTAAGATTGGTTAAAGTCCCGCAACGAGCGCA 1080  
Qy 1078 ACCCTTGTCAATTAATTCGATCACTTTGGTTGGGCACTTTAATGACGCGGTGACAA 1137  
Db 1081 ACCCTTGTCAATTAATTCGATCACTTTGGTTGGGCACTTTAATGACGCGGTGACAA 1140  
Qy 1138 CCGGAGGAGGTTGGGAGTGAAGTCAAGTCTCATGGCCCTTATGGTGGGCTTTCACAG 1197  
Db 1141 CCGGAGGAGGTTGGGAGTGAAGTCAAGTCTCATGGCCCTTATGGTGGGCTTTCACAG 1200  
Qy 1198 TAATACATGCGCGCTGACAGAGGTTGCCAACCCCGGAGGAGCTAATCTCAGAAAGC 1257  
Db 1201 TAATACATGCGCGCTGACAGAGGTTGCCAACCCCGGAGGAGCTAATCTCAGAAAGC 1260  
Qy 1258 GCGTCTGATGTCGGAATGGAAGTCTGCAACTGCTCCGTGAATCCGGAATCCGTATGAAT 1317  
Db 1261 GCGTCTGATGTCGGAATGGAAGTCTGCAACTGCTCCGTGAATCCGGAATCCGTATGAAT 1320  
Qy 1318 CCGGAGTACAGATGTCGCGGTGAATAGCTCCCGGCTTGTACACACGCGCGTCAAC 1377  
Db 1321 CCGGAGTACAGATGTCGCGGTGAATAGCTCCCGGCTTGTACACACGCGCGTCAAC 1380  
Qy 1378 CATGGAGTGGTTCACAGAGCAGTGTCTAACCGTAAAGAGGCGCTTGCACAG 1437  
Db 1381 CATGGAGTGGTTCACAGAGCAGTGTCTAACCGTAAAGAGGCGCTTGCACAG 1440  
Qy 1438 TGAGATTGATGACTGGGGTG 1457  
Db 1441 TGAGATTGATGACTGGGGTG 1460

RESULT 64  
AY123788 1497 bp DNA linear BCT 23-SEP-2003  
LOCUS Nitrosospiira sp. Nsp41 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY123788  
ACCESSION AY123788  
VERSION AY123788.1 GI:24474417  
KEYWORDS  
SOURCE Nitrosospiira sp. Nsp41  
ORGANISM Nitrosospiira sp. Nsp41  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospiira.  
REFERENCE 1 (bases 1 to 1497)  
Purkhof, U., Wagner, M., Timmermann, G., Pommerening-Rosser, A. and  
Koops, H. P.  
JOURNAL 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
FUBMED ammonia-oxidizing isolates: extension of the dataset and proposal  
AUTHORS of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
13130037  
2 (bases 1 to 1497)  
Purkhof, U., Timmermann, G., Koops, H. P., Pommerening-Rosser, A. and  
Wagner, M.

TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1..1497  
/organism="Nitrosospiira sp. Nsp41"  
/mol\_type="genomic DNA"  
/strain="Nsp41"  
/db\_xref="taxon:200115"  
<1..>1497  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 88.3%; Score 1286; DB 1; Length 1497;  
Best Local Similarity 93.5%; Pred. No. 3 7e-13;  
Matches 1365; Conservative 0; Mismatches 90; Indels 5; Gaps 2;  
Qy 1 ATTGAACGCTGCGGCAATGCTTTACATGCAATGCAACGCGACACGGAATCTTGCAT 60  
Db 3 ATTGAACGCTGCGGCAATGCTTTACATGCAATGCAACGCGACACGCGGGCAAC-- 60  
Qy 61 CTGTGCGAGTGGCGGACGCGGTGATTAATGATCGGAACGTAATCAGAAAGGGGGGTA 120  
Db 61 CTGTGCGAGTGGCGGACGCGGTGATTAATGATCGGAACGTAATCAGAAAGGGGGGTA 120  
Qy 121 AGCGATCGAAAGATGCTTAATCCGCATATATCTTAAGAGAGAAAGCGGGGATCGAA 180  
Db 121 AGCGATCGAAAGATGCTTAATCCGCATATATCTTAAGAGAGAAAGCGGGGATCGAA 180  
Qy 181 GACCTTGGCTTTGGAGCGCGCATGCTGTAGTATGCTAGTTGTTGGGTTAAAGGCTTAC 240  
Db 181 GACCTTGGCTTTGGAGCGCGCATGCTGTAGTATGCTAGTTGTTGGGTTAAAGGCTTAC 240  
Qy 241 CAAGCGACATCAATGATTGTTGTTGAGAGAGACGACACCACTGGGACTGAGACAG 300  
Db 241 CAAGCGACATCAATGATTGTTGTTGAGAGAGACGACACCACTGGGACTGAGACAG 300  
Qy 301 GCCCAGACTCTTACGCGGAGGACGACATGGGGAAATTTTGAACAATGGGGGAAACCTTATC 360  
Db 301 GCCCAGACTCTTACGCGGAGGACGACATGGGGAAATTTTGAACAATGGGGGAAACCTTATC 360  
Qy 361 CAGCAATGCGGGTGAAGAGGCTTGGGTTGTAAGCTCTTTCAGTGAAGAGA 420  
Db 361 CAGCAATGCGGGTGAAGAGGCTTGGGTTGTAAGCTCTTTCAGTGAAGAGA 420  
Qy 421 AAAGTTACGTTAATTAATCTGATCATGACGTTATCGACAGAAAGACACCGGCTAAC 480  
Db 421 AAAGTTACGTTAATTAATCTGATCATGACGTTATCGACAGAAAGACACCGGCTAAC 480  
Qy 481 TACGTCCAGACGCGCGGTAAATAGTGAAGGTCGAAAGCGTTAATCGGAATTACTGGCGT 540  
Db 481 TACGTCCAGACGCGCGGTAAATAGTGAAGGTCGAAAGCGTTAATCGGAATTACTGGCGT 540  
Qy 541 AAAGGTGGCGGAGCGGCTTTGTAAGTCAATGTAATCCCGGCTTAACTGGGAAT 600  
Db 541 AAAGGTGGCGGAGCGGCTTTGTAAGTCAATGTAATCCCGGCTTAACTGGGAAT 600  
Qy 601 TGCGTTGAAACTCAAGAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTAGCAG 660  
Db 601 TGCGTTGAAACTCAAGAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTAGCAG 660  
Qy 661 TGAATGCTGATGATTAAGAGAAACATGATGCGGACGACCTCTGGTTAACTACT 720  
Db 661 TGAATGCTGATGATTAAGAGAAACATGATGCGGACGACCTCTGGTTAACTACT 720  
Qy 721 GAGCTTCAGCAAGAAAGGTGGGAGCAAAAGAGATTAGATACCTGTAGTCCAGCC 780  
Db 721 GAGCTTCAGCAAGAAAGGTGGGAGCAAAAGAGATTAGATACCTGTAGTCCAGCC 780  
Qy 781 CTAAGCATGTCACTAGTTGTTGGCTTATTAATGCTTGTGTAAGAGCTAACGCGTGA 840  
Db 781 CTAAGCATGTCACTAGTTGTTGGCTTATTAATGCTTGTGTAAGAGCTAACGCGTGA 840



QY	841	AGTTGACCCGCTGGGGAGTACGGTCCGAAGTTAAACTCAAGAATTTGACGGGACC	900
Db	841	AGTTGACCGGCTGGGGAGTACGGTCCGAAGTTAAACTCAAGAATTTGACGGGACC	900
QY	901	GCACAAGCGGTGGATTATGTGGAATTAAATTCGATGCAAGCGCAAAAACCTTACTACCTT	960
Db	901	GCACAAGCGGTGGATTATGTGGAATTAAATTCGATGCAAGCGCAAAAACCTTACTACCTT	960
QY	961	GACATGTAGCGAAATTTTCTAGAGATAGATTAGTGC--TTCCGGAAACGCTAACAGGTG	1017
Db	961	GACATGTAGCGAAAGCCCGCGAGAGGTGGGTGTGCCGAAGGAGCGGTAAACAAGGTG	1020
QY	1018	CTGCATGGCTGTCTGCTCACTGTGTCTGTAAGATGTGGGTTAAATGTCCTCCGACAAGCGCA	1077
Db	1021	CTGCATGGCTGTCTGCTCACTGTGTCTGTAAGATGTGGGTTAAATGTCCTCCGACAAGCGCA	1080
QY	1078	ACCGTGTGATTAATTTGCTATCTTTGGTTGGGCACTTTAATGAGACTGCCGTGACAA	1137
Db	1081	ACCGTGTGATTAATTTGCTATCTTTGGTTGGGCACTTTAATGAGACTGCCGTGACAA	1140
QY	1138	CCGGAGGAAGGTGGGGATGACGTCAAGTCTCATATGCCCCCTTAATGGGTAGGGCTTACACG	1197
Db	1141	CCGGAGGAAGGTGGGGATGACGTCAAGTCTCATATGCCCCCTTAATGGGTAGGGCTTACACG	1200
QY	1198	TAAATCAATGCGCGCTTACAGAGGTTGCCAACCCCGAGGGGAGCTATCTCAGAAAGC	1257
Db	1201	TAAATCAATGCGCGCTTACAGAGGTTGCCAACCCCGAGGGGAGCTATCTCAGAAAGC	1260
QY	1258	GGGTCTGTATGTCGGATCGGAGTCTGCACCTGACCTCGGTGAAGTCCGATGCTAGTAAT	1317
Db	1261	GGGTCTGTATGTCGGATCGGAGTCTGCACCTGACCTCGGTGAAGTCCGATGCTAGTAAT	1320
QY	1318	CGCGGATCAGATGTCCGGTGAATACCTTCCGGGCTTGTACACACCGCCCGTCAAC	1377
Db	1321	CGCGGATCAGATGTCCGGTGAATACCTTCCGGGCTTGTGTACACACCGCCCGTCAAC	1380
QY	1378	CATGGAGTGGGTTTACCAAGACAGGTAGTCTAACCTTAAGAGAGGCGCTTGCCACGG	1437
Db	1381	CATGGAGTGGGTTTACCAAGACAGGTAGTCTAACCGCAAGAGGCGCGCTTGCCACGG	1440
QY	1438	TGAGATTGATGATCGGGGTG	1457
Db	1441	TGAGATTGATGATCGGGGTG	1460
RESULT 65			
AF386755			
LOCUS	AF386755	1469 bp	DNA linear ENV 10-MAY-2004
DEFINITION	Nitrosospirila sp. R5c20 16S ribosomal RNA gene, partial sequence.		
ACCESSION	AF386755		
VERSION	AF386755.1	GI:17864832	
KEYWORDS	ENV.		
SOURCE	Nitrosospirila sp. R5c20		
ORGANISM	Nitrosospirila sp. R5c20		
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosospirila; environmental samples.		
AUTHORS	1 (bases 1 to 1469)		
TITLE	Burrell, P. C., Phalen, C.M. and Hovanec, T.A.		
JOURNAL	Identification of bacteria responsible for ammonia oxidation in freshwater aquaria		
PUBMED	Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)		
AUTHORS	2 (bases 1 to 1469)		
TITLE	Burrell, P. C., Phalen, C.M. and Hovanec, T.A.		
JOURNAL	Direct Submission		
FEATURES	Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria Group, 6100 Condon Dr, Moorpark, CA 93021, USA		
source	Location/Qualifiers		
	1. 1469		
	/organism="Nitrosospirila sp. R5c20"		
	/mol_type="genomic DNA"		
	/isolation_source="ammonia-oxidizing bacteria isolated from freshwater aquaria enrichments"		

Query Match	Best Local Similarity	Matches 1362; Conservative	88.2%; 93.7%; 0;	Score 1284.8; Pred. No. 4e-13; Mismatches 87;	DB 3; Length 1469; Indels 5; Gaps 2
1	ATTGAACGCTGCGCGCATCTTATACATCGAATGCAACGGCAGCAGCATGCTTGAT	60			
18	ATTGAACGCTGCGCGCATCTTATACATCGAATGCAACGGCAGCAGCATGCGGCGAAC	75			
61	CTGTGTGCGAGATGCGCGAAGCGGTGATATCATTCGAAACGTATCCAGAAAGGGGGGTA	120			
76	CTGTGTGCGAGATGCGCGAAGCGGTGATATCATTCGAAACGTATCCAGATGCGGGGTA	135			
121	AGCGATCGAAAGATGTGCTTAATACCGCATATATCTTAAAGAGGAAACACAGGATCGAA	180			
136	ACCGACCGAAAGATGTGCTTAATACCGCATATATCTTAAAGAGGAAACACAGGATCGCA	195			
181	GACCTTGCGCTTTTGAGAGCGCGCATGCTATAGTATAGTATGCTGCGGTGAGGAGCTTAC	240			
196	GACCTTGCGCTTTTGAGAGCGCGCATGCTATAGTATAGTATGCTGCGGTGAGGAGCTTAC	255			
241	CAAGCGAGCATCAGTATGTTGCTGAGAGAGACACACGACCATGCGAATGAGACAG	300			
256	CAAGCGAGCATCAGTATGTTGCTGAGAGAGACACACGACCATGCGAATGAGACAG	315			
301	GCCCAAGCTCTTACCGGAGGACGACGTGGGAAATTTTGACAAATGCGCGCAAGCTGATC	360			
316	GCCCAAGCTCTTACCGGAGGACGACGTGGGAAATTTTGACAAATGCGCGCAAGCTGATC	375			
361	CACCAATGCGCGGTGAGTGAAGAAGCGCTTGGGTTGAAAGCTCTTACGTCGAGAGA	420			
376	CACCAATGCGCGGTGAGTGAAGAAGCGCTTGGGTTGAAAGCTCTTACGTCGAGAGA	435			
421	AAAGGTTACGCTAAATATCTGATCTCATGACGGTATCGACAGAAAGACACCGGCTAAC	480			
436	AAAGGTTACGCTAAATATCTGATCTCATGACGGTATCGACAGAAAGACACCGGCTAAC	495			
481	TACGTCCACAGACCGCGGTAAATACGTAGGGGTGACAGGCTTAATCGGAATTACTGGGCGT	540			
496	TACGTCCACAGACCGCGGTAAATACGTAGGGGTGACAGGCTTAATCGGAATTACTGGGCGT	555			
541	AAAGGTTGCGCGCGGCTTGTAGTACAGATGGAATCCCGGGGCTTAAACCTGGGAAT	600			
556	AAAGGTTGCGCGCGGCTTGTAGTACAGATGGAATCCCGGGGCTTAAACCTGGGAAT	615			
601	TGCGTTTGAACCTACAGAGCTAGAGTGTGAGAGAGGAGGTGAAATTCATGTGTAGCAG	660			
616	TGCGTTTGAACCTACAGAGCTAGAGTGTGAGAGAGGAGGTGAAATTCATGTGTAGCAG	675			
661	TGAATATCGTAAGATATGAAAGAAATCATGATGCGAAAGCAGCTCTCTGGGTTAACACT	720			
676	TGAATATCGTAAGATATGAAAGAAATCATGATGCGAAAGCAGCTCTCTGGGTTAACACT	735			
721	GACGCTCATGACGAAAGCGTGGGAGGAAACAGATTAAGTAACTCCGTGATAGTCCAGCC	780			
736	GACGCTCATGACGAAAGCGTGGGAGGAAACAGATTAAGTAACTCCGTGATAGTCCAGCC	795			
781	CTAAAGCATGTCAACTAGTGTGTGGGCTTTATAGGCTTGGTATACGAAGCTAACGCGTGA	840			
796	CTAAAGCATGTCAACTAGTGTGTGGGCTTTATAGGCTTGGTATACGAAGCTAACGCGTGA	855			
841	AGTTGACCGCTGCGGAGTACCGTGCAGAAATTAATCTCAAGAAATTTGACCGGAGACC	900			
856	AGTTGACCGCTGCGGAGTACCGTGCAGAAATTAATCTCAAGAAATTTGACCGGAGACC	915			
901	GCAACAAGCGGTGATATATGATTAATTCGATGCAACGCAAAACCTTAACCTACCTT	960			



Db 916 GCACAAGCGGTGATTTGATGATTAATTCAGTACGCAAGCGGAAAACTTACCTTACCTT 975  
Qy 961 GACATGTACGGAATTTTCTAGAGATAGATTAAGTGC---TTCCGGAAACGCTTAACACAGGTG 1017  
Db 976 GACATGTACGGAATTTTCTAGAGATAGATTAAGTGC---TTCCGGAAACGCTTAACACAGGTG 1035  
Qy 1018 CTGCATGAGCTGTCTGTCAGCTGTCTGTCAGATGTTGGTTAAAGTCCCGCAACGAGCGCA 1077  
Db 1036 CTGCATGAGCTGTCTGTCAGCTGTCTGTCAGATGTTGGTTAAAGTCCCGCAACGAGCGCA 1095  
Qy 1078 ACCCTGTCAATTAATGTCATCACTTGTGTCGATTTAAATGAGCTCCGCTGACAA 1137  
Db 1096 ACCCTGTCAATTAATGTCATCACTTGTGTCGATTTAAATGAGCTCCGCTGACAA 1155  
Qy 1138 CCGGAGGAGGTGGGGATGACGTCAGTCTCTCATGAGCCCTTAATGAGTGGGCTTCAACAG 1197  
Db 1156 CCGGAGGAGGTGGGGATGACGTCAGTCTCTCATGAGCCCTTAATGAGTGGGCTTCAACAG 1215  
Qy 1198 TAATCAATGAGCGCTGACAGAGGTTGCCAACCCCGGAGGGGAGCTTAATCTCAGAAAGC 1257  
Db 1216 TAATCAATGAGCGCTGACAGAGGTTGCCAACCCCGGAGGGGAGCTTAATCTCAGAAAGC 1275  
Qy 1258 GCGTGTAGTCCGAGTCCGAGTCTGCACTGCACTCCGTAAGTCCGTAATCTGTAAT 1317  
Db 1276 GCGTGTAGTCCGAGTCCGAGTCTGCACTGCACTCCGTAAGTCCGTAATCTGTAAT 1335  
Qy 1318 CCGGATGACAGTGTGCGGTGAAATAGTTCCCGGGCTTGTACACACCGCCGCTACAC 1377  
Db 1336 CCGGATGACAGTGTGCGGTGAAATAGTTCCCGGGCTTGTACACACCGCCGCTACAC 1395  
Qy 1378 CATGAGAGTGGTGTTCACAGAGCAGTAGTCTAAACCTTAAGAGAGCGCTTGCACAG 1437  
Db 1396 CATGAGAGTGGTGTTCACAGAGCAGTAGTCTAAACCTTAAGAGAGCGCTTGCACAG 1455  
Qy 1438 TGAGATTCACTACT 1451  
Db 1456 TGAGATTCACTACT 1469

RESULT 66  
AF386754 1486 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosospira sp. Bf16c46 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386754  
ACCESSION AF386754  
VERSION AF386754.1 GI:17864831  
KEYWORDS ENV.  
SOURCE Nitrosospira sp. Bf16c46  
ORGANISM Nitrosospira sp. Bf16c46  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira; environmental samples.  
REFERENCE 1 (bases 1 to 1486)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 1172936  
REFERENCE 2 (bases 1 to 1486)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 1000 Condon Dr, Moorpark, CA 93021, USA

FEATURES  
source  
1..1486  
/organism="Nitrosospira sp. Bf16c46"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichments"  
/db\_xref="taxon:180513"  
/clone="Bf16c46"  
/environmental\_sample  
<1..>1486  
/product="16S ribosomal RNA"

ORIGIN  
rRNA

Query Match 88.2%; Score 1284.4; DB 3; Length 1486;  
Best Local Similarity 93.4%; Pred. No. 4e-13; 91; Indels 5; Gaps 2;  
Matches 1364; Conservative 0; Mismatches 91;  
Qy 1 ATGGAACGCTGGCGCATGCTTTACATGCAAGTCGACGCGACGATGCTTGCAT 60  
Db 16 ATGGAACGCTGGCGCATGCTTTACATGCAAGTCGACGCGACGCGCGCAAC-- 73  
Qy 61 CTGTGGGAGTGGCGGACGGGTGATATGATGGAACCTATCTCAAGAGGGGGGTA 120  
Db 74 CTGTGGGAGTGGCGGACGGGTGATATCATGGAACCTATCTCGAGTGGGGGTA 133  
Qy 121 AGCGATCGAAGATGCTTAATACCGCATATCTTAAGAGGAAAGCAGGGATCGAA 180  
Db 134 AGCGATCGAAGATGCTTAATACCGCATATCTTAAGAGGAAAGCAGGGATCGAA 193  
Qy 181 GACCTTGGCTTTTGAAGCGCGCATGCTGATTAAGTGTGTTGGGTAAAGCTTAC 240  
Db 194 GACCTTGGCTTTTGAAGCGCGCATGCTGATTAAGTGTGTTGGGTAAAGCTTAC 253  
Qy 241 CAAGCGACGATCAGTAGTTGTTGAGAGACGACCACTGGGACTGAGACAG 300  
Db 254 CAAGCGACGATCAGTAGTTGTTGAGAGACGACCACTGGGACTGAGACAG 313  
Qy 301 GCCCAGCTCTTACCGGAGGCGACAGTGGGGAATTTTGAACATGGGCGCAAGCTGATC 360  
Db 314 GCCCAGCTCTTACCGGAGGCGACAGTGGGGAATTTTGAACATGGGCGCAAGCTGATC 373  
Qy 361 CAGCAATGCGGGGTGAGTGAAGAGGCTTCCGGTTGTAAAGCTCTTTCAGTCAGAAAG 420  
Db 374 CAGCAATGCGGGGTGAGTGAAGAGGCTTCCGGTTGTAAAGCTCTTTCAGTCAGAAAG 433  
Qy 421 AAAGGTTACGTTAAATTAATCTGATCATGACGATTCGACAGAAAGACACCGGCTAAC 480  
Db 434 AAAGGTTACGTTAAATTAATCTGATCATGACGATTCGACAGAAAGACACCGGCTAAC 493  
Qy 481 TACGTCCACAGCCCGGCTAATGCTAGAGGTGCAAGCCTTAATCGGAATTTCTGGGCGT 540  
Db 494 TACGTCCACAGCCCGGCTAATGCTAGAGGTGCAAGCCTTAATCGGAATTTCTGGGCGT 553  
Qy 541 AAAGGTTGCGGCGGCGGCTTTGTAAGTCAGATGTGAATCCCGGCTTAACCTGGGAAT 600  
Db 554 AAAGGTTGCGGCGGCGGCTTTGTAAGTCAGATGTGAATCCCGGCTTAACCTGGGAAT 613  
Qy 601 TCGGTTGAAACTACAGGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTACAG 660  
Db 614 TCGGTTGAAACTACAGGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTACAG 673  
Qy 661 TGAATGCGTAGATATGGAAGACATGATGCGGAGGAGGAGGCTTCCGGTTAACT 720  
Db 674 TGAATGCGTAGATATGGAAGACATGATGCGGAGGAGGAGGCTTCCGGTTAACT 733  
Qy 721 GAGGCTCAGCAGAAAGGTGGGAGCAAAAGATTAATGCTCTGTAAGTCCAGCGC 780  
Db 734 GAGGCTCAGCAGAAAGGTGGGAGCAAAAGATTAATGCTCTGTAAGTCCAGCGC 793  
Qy 781 CTAACGATGCACTAGTGTGTTGGCTTATTAAGCTTGTGAACGAAAGCTTAACGCGCTGA 840  
Db 794 CTAACGATGCACTAGTGTGTTGGCTTATTAAGCTTGTGAACGAAAGCTTAACGCGCTGA 853  
Qy 841 AGTTGACCGCTGGGAGTACGCTGCCAAGATTAACTCAAGGAATTAACGCGGAGCC 900  
Db 854 AGTTGACCGCTGGGAGTACGCTGCCAAGATTAACTCAAGGAATTAACGCGGAGCC 913  
Qy 901 GCACAAGCGGTGATATGAGTAAATTCGATGCAACGCGGAAACCTTAACCTACCTT 960  
Db 914 GCACAAGCGGTGATATGAGTAAATTCGATGCAACGCGGAAACCTTAACCTACCTT 973  
Qy 961 GACATGTACGGAATTTTCTAGAGATAGTATGTC---TTCCGGAAACGCTTAACACAGGTG 1017  
Db 974 GACATGTACGGAATTTTCTAGAGATAGTATGTC---TTCCGGAAACGCGGTAAACACAGGTG 1033

QY 1018 CTGCATGGCTGTCAGCTGTCGTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1077  
DB 1034 CTGCATGGCTGTCAGCTGTCGTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1093  
QY 1078 ACCCTTGTCAATTAATTGCCATATTGGTTGGCACTTTAATAGACTGCCGGTACAA 1137  
DB 1094 ACCCTTGTCAATTAATTGCCATATTGGTTGGCACTTTAATAGACTGCCGGTACAA 1153  
QY 1138 CCGGAGGAAGGTGGGAGTGAACGTCTCATGGCCCTTAATGGGTAGGCTTCAACG 1197  
DB 1154 CCGGAGGAAGGTGGGAGTGAACGTCTCATGGCCCTTAATGGGTAGGCTTCAACG 1213  
QY 1198 TAAATCAATGGCGCTTACAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAGAAAGC 1257  
DB 1214 TAAATCAATGGCGCTTACAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAGAAAGC 1273  
QY 1258 GCGTGTGATCCGAGTCGAGTCTGCAACTCGACTCCGTAAGTGGAAATCGCTAGTAAT 1317  
DB 1274 GCGTGTGATCCGAGTCGAGTCTGCAACTCGACTCCGTAAGTGGAAATCGCTAGTAAT 1333  
QY 1318 CCGGATCAAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTAACAACCCCGCTCAAC 1377  
DB 1334 CCGGATCAAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTAACAACCCCGCTCAAC 1393  
QY 1378 CATGGAGTGGGTTCAACGAGAGAGGTAGTCTAACCGTAAAGAGGGGCTTGCACAG 1437  
DB 1394 CATGGAGTGGGTTCAACGAGAGAGGTAGTCTAACCGTAAAGAGGGGCTTGCACAG 1453  
QY 1438 TGAGATTCACTGACTGGGGTG 1457  
DB 1454 TGAGATTCACTGACTGGGGTG 1473

RESULT 67  
AJ621030 1459 bp DNA linear BCT 08-JUN-2004  
LOCUS Nitrosospiira sp. Is148 16S rRNA gene, isolate Is148.  
DEFINITION AJ621030  
ACCESSION AJ621030.1 GI:40994849  
VERSION 16S ribosomal RNA; 16S rRNA gene.  
KEYWORDS Nitrosospiira sp. Is148  
SOURCE Nitrosospiira sp. Is148  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosospiira.

REFERENCE  
1. Nielsen, S., Baer Gyllissen, M.J., Revsbech, N., Leandbroek, H.J. and Bollmann, A.  
Influence of heterotrophic bacteria on the activity and growth behavior of ammonia-oxidizing bacteria  
Unpublished 2 (bases 1 to 1459)

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Bollmann, A.  
Direct Submision  
Submitted (15-JAN-2004) Bollmann A., Department for Microbial Ecology, NIOO-KNAW Centre for Limnology, Rijksstraatweg 6, 3631 AC Nieuwersluis, NETHERLANDS  
Location/Qualifiers

FEATURES  
source  
1. 1459  
/organism="Nitrosospiira sp. Is148"  
/mol\_type="genomic DNA"  
/isolate="Is148"  
/isolation\_source="fresh water sediment"  
/db\_xref="taxon:261296"  
/country="Belgium:Apeldoorn"  
1. 1459  
/gene="16S rRNA"  
1. 1459  
/gene="16S rRNA"  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 88.1%; Score 1283.8; DB 1; Length 1459;  
Best Local Similarity 94.2%; Pred. No. 4.2e-13;

Matches 1367; Conservative 0; Mismatches 77; Indels 7; Gaps 3;  
QY 1 ATTTGAACCTTGGCGGCAATGCTTTACATGCAATGCAAGGCGACGAGTGTTCAT 60  
DB 13 ATTTGAACCTTGGCGGCAATGCTTTACATGCAATGCAAGGCGACGAGTGTTCAT 69  
QY 61 CTGGTGGGAGTGGCGGAGCGGTGATTAATGATTCGCAACCTATTCAGAAAGGCGGTA 120  
DB 70 CTGGTGGGAGTGGCGGAGCGGTGATTAATGATTCGCAACCTATTCAGAAAGGCGGTA 129  
QY 121 ACGCATCAAAAGATGTGTAATACCGCATATCTCTAAGAGAAAGCAGGGATCGAAA 180  
DB 130 ACGCATCAAAAGATGTGTAATACCGCATATCTCTAAGAGAAAGCAGGGATCGAAA 189  
QY 181 GACCTTGGCTTTTGGAGCGGCGCATGTCTGATTAAGTGTGGGTAAAGGCTTAC 240  
DB 190 GACCTTGGCTTTTGGAGCGGCGCATGTCTGATTAAGTGTGGGTAAAGGCTTAC 249  
QY 241 CAAGCGCAGATCAAGTATGTTGTCAGAGAGACGACGACCACTGGGACTGAGACAG 300  
DB 250 CAAGCTTCAATCAAGTATGTTGTCAGAGAGACGACGACCACTGGGACTGAGACAG 309  
QY 301 GCCCAACTCTTACGGGAGGAGCAGCATGGGAAATTTTGGACATGGGCGCAAGCTATC 360  
DB 310 GCCCAACTCTTACGGGAGGAGCAGCATGGGAAATTTTGGACATGGGCGCAAGCTATC 369  
QY 361 CAGCAATGCCCGTGAAGAGAAAGGCTTGGGGTGTAAAGCTCTTTCAGTTCAGAAAG 420  
DB 370 CAGCAATGCCCGTGAAGAGAAAGGCTTGGGGTGTAAAGCTCTTTCAGTTCAGAAAG 429  
QY 421 AAAGTTACGGTAAATATCTGTACTCATGACGGTATCGAAGAAAGACCGGCTAAC 480  
DB 430 AAAGTTACGGTAAATATCTGTACTCATGACGGTATCGAAGAAAGACCGGCTAAC 489  
QY 481 TACGTCCAGCAGACCGCGGTAATACGTAGGTCGACGCTTAATTCGAAATTCGGGGCT 540  
DB 490 TACGTCCAGCAGACCGCGGTAATACGTAGGTCGACGCTTAATTCGAAATTCGGGGCT 549  
QY 541 AAAGGTGGCGAGGCGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAAAT 600  
DB 550 AAAGGTGGCGAGGCGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAAAT 609  
QY 601 TGGCTTTGAATCAAAAGCTAGAGTGTGCGACGAGGAGGTGAATTCATGTGTAGCAG 660  
DB 610 TGGCTTTGAATCAAAAGCTAGAGTGTGCGACGAGGAGGTGAATTCATGTGTAGCAG 669  
QY 661 TGAATGTGTAAGATATGAAAGACATCATGTGCGAAGGACGCTCTGGGTTAACCT 720  
DB 670 TGAATGTGTAAGATATGAAAGACATCATGTGCGAAGGACGCTCTGGGTTAACCT 729  
QY 721 GACGCTATGCAACGAAAGCGTGGGAGGAAACAGGATTAATCCCTGTAGTCCAGCGC 780  
DB 730 GACGCTATGCAACGAAAGCGTGGGAGGAAACAGGATTAATCCCTGTAGTCCAGCGC 789  
QY 781 CTAACGATGTCATAGTGTGCGCTTAAATGAGCTTGGTAAAGAACTTAACGCGTGA 840  
DB 790 CTAACGATGTCATAGTGTGCGCTTAAATGAGCTTGGTAAAGAACTTAACGCGTGA 849  
QY 841 AGTTGACCGCTGGGAGTACGCTGCGCAAGTTAAATCTCAAGAAATTGACGGGAGACC 900  
DB 850 AGTTGACCGCTGGGAGTACGCTGCGCAAGTTAAATCTCAAGAAATTGACGGGAGACC 908  
QY 901 GCACAAAGGCGGATTAATGAGTAAATTCGATGCAAGCGGAAACCTTACCTACCTT 960  
DB 909 GCACAAAGGCGGATTAATGAGTAAATTCGATGCAAGCGGAAACCTTACCTACCTT 968  
QY 961 GACATGTAGCAATTTTCTAGAGATAGATTAGTGC---TTGCGGAACGCTAACAGGTTG 1017  
DB 969 GACATGTAGCAATTTTCTAGAGATAGATTAGTGC---TTGCGGAACGCTAACAGGTTG 1028  
QY 1018 CTGCATGGCTGTCAGCTGTCGTGAGATGTTGGGTTAAGTCCGCAACGAGCGCA 1077  
DB 1029 CTGCATGGCTGTCAGCTGTCGTGAGATGTTGGGTTAAGTCCGCAACGAGCGCA 1088

Qy	1078	ACCTTGTCAATTAATTGTCATCATTTGGTTGGGCATTTAATGAGACTGCGCGTGACAA	1137
Db	1089	ACCTTGTCAATTAATTGTCATCATTTGAGGCACTTTAATGAACTGCCGCTGACAA	1148
Qy	1138	CCGAGAGAGGTGGGGATGACGTCMAAGTCCTCATGCGCCCTTAATGGTAGGGCTTCAACG	1197
Db	1149	CCGAGAGAGGTGGGGATGACGTCMAAGTCCTCATGCGCCCTTAATGGTAGGGCTTCAACG	1208
Qy	1198	TAATPACAATGGCGCGGTACAGAGGGTTGGCAACCCGCGAGGGGGAGCTMAATCTCAGAAAGC	1257
Db	1209	TAATPACAATGGCGCGGTACAGAGGGTTGGCAACCCGCGAGGGGGAGCTMAATCTCAGAAAGC	1268
Qy	1258	GCGTCGTAAGTCGCGATGGGAATGCGCAACTGCACTCGGTGAAGTCGGAATCGCTAATAT	1317
Db	1269	GCGTCGTAAGTCGCGATGGGAATGCGCAACTGCACTCGGTGAAGTCGGAATCGCTAATAT	1328
Qy	1318	CGCGGATCAGCATGTCGCGGTGAATPACGTTCCGGGCTTGTGACACACCGCGCTGCACAC	1377
Db	1329	CGCGGATCAGCATGTCGCGGTGAATPACGTTCCGGGCTTGTGACACACCGCGCGTGCACAC	1388
Qy	1378	CATGGAGATGGGCTTACCAAGACAGGTATCTMAACCGTAAGAGGGCGCTTGCCACGG	1437
Db	1389	CATGGAGATGGGCTTACCAAGACAGGTATCTMAACCGCAAGAGGGCGCTTGCCACGG	1448
Qy	1438	TGAGATTGATG	1448
Db	1449	TGAGATTGATG	1459
RESULT 68			
LOCUS	AY856375	1458 bp	DNA linear BCT 15-JAN-2005
DEFINITION	Nitrosospiira sp. HBN8222A 16S ribosomal RNA gene, partial sequence.		
ACCESSION	AY856375		
VERSION	AY856375.1	GI:57338938	
KEYWORDS			
SOURCE			
ORGANISM	Nitrosospiira sp. HBN8222A		
	Nitrosospiira sp. HBN8222A		
	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;		
	Nitrosomonadaceae; Nitrosospiira.		
	1 (bases 1 to 1458)		
REFERENCE	Takahashi, R., Yabe, R. and Tokuyama, T.		
AUTHORS	Submitted (14-DEC-2004) Agricultural and Biological Chemistry,		
TITLE	Direct Submission		
JOURNAL	Submitted (14-DEC-2004) Agricultural and Biological Chemistry,		
	College of Bioresource Sciences, Nihon University, 1866 Kameiyo,		
	Fujiisawa, Kanagawa 252-8510, Japan		
FEATURES			
source	Location/Qualifiers		
	1..1458		
	/organism="Nitrosospiira sp. HBN8222A"		
	/mol_type="genomic DNA"		
	/strain="HBN8222A"		
	/db_xref="taxon:308063"		
	/note="similar to Nitrosospiira (Nitrosolobus) multiformis"		
	<1..>1458		
	/product="16S ribosomal RNA"		
rRNA			
ORIGIN			
Query Match	88.0%;	Score 1282.8;	DB 1; Length 1458;
Best Local Similarity	93.4%;	Pred. No. 4.3e-13;	
Matches 1363; Conservative	0;	Mismatches 92;	Indels 5; Gaps 2;
Qy	1	ATTGAACGCTGGCGCGCATCTTTACACATGCAAGTCGACCGCAGACACGAGTGTTCGAT	60
Db	1	ATTGAACGCTGGCGCGCATCTTTACACATGCAAGTCGACCGCAGACACGAGTGTTCGAT	58
Qy	61	CTGTGTGGCAGTGGCGGACCGGTTGAGTAATGCATCGAAACGTATCCAGAAAGGGGGGTA	120
Db	59	CTGTGTGGCAGTGGCGGACCGGTTGAGTAATGCATCGAAACGTATCCTTAAGTGGGGAGTA	118
Qy	121	ACGCATCGAAAGATGCTAATPACCGCATATCTCTPAAAGAGAAAGACGAGGATTCGAA	180
Db	119	ACGCATCGAAAGATGCTAATPACCGCATATCTCTGAGAGAAAGACGAGGATTCGCA	178

QY	181	GACCTTGCCTTTTGGAGCGGCCGAGTGTCTGATTTACTGTGTGTTGGGTAAAGGCGTCA	240
Db	179	GACCTTGCCTTTTGGAGCGGCCGAGTGTCTGATTTACTGTGTGTTGGGTAAAGGCGTCA	238
QY	241	CAAGGCGCAGCATAGTATGTGTGTCTAGAGAGACGACCGACACACTGTGGACTGAGACAG	300
Db	239	CAGGCGCTCAATCAATGATGCTGTCTGAGAGAGACGACCGACACACTGTGGACTGAGACAG	298
QY	301	GCCGAGACTCTTAACGGGAGGCGACAGTGGGAAATTTTGCACATTTGGCGCGACGCTGATC	360
Db	299	GCCGAGACTCTCTAACGGGAGGCGACAGTGGGAAATTTTGCACATTTGGCGCGAAACCTGTATC	358
QY	361	CAGCAATGCCCGCTGAGTGAAGAAGCGCTTCGGGTGTGAAGCTCTTTCAAGCGGAAACGA	420
Db	359	CAGCAATGCCCGCTGAGTGAAGAAGCGCTTCGGGTGTGAAGCTCTTTCAAGCGGAAACGA	418
QY	421	AAAGGTAAACGGTAAATTAATGATGATCATATGACGGTATCGACAGAAAGACCGCGTAAAC	480
Db	419	AAAGGTAAACGGTAAATTAATGATGATCATATGACGGTATCGACAGAAAGACCGCGGTACC	478
QY	481	TACGTGCAGACGCCGCGGTAAATCGTAGGGTGCAGACGTTAAATCGGAATTAATCTGGCGCT	540
Db	479	TACGTGCAGACGCCGCGGTAAATCGTAGGGTGCAGACGTTAAATCGGAATTAATCTGGCGCT	538
QY	541	AAAGGGTGCAGACGCCGCTTTGTAAATGTCAGATGTGAATATCCCGGGCTTAAACCTGGGAAT	600
Db	539	AAAGCGTGCAGACGCCGCTTTGTAAATGTCAGATGTGAATATCCCGGGCTTAAACCTGGGAAC	598
QY	601	TGCGTTTGAATCTAACAGGCTAGATGTGGCAGAGGAGGTGGAAATTCATATGTTAGACAG	660
Db	599	TGCGTTTGAATCTAACAGGCTAGATGTGGCAGAGGAGGTGGAAATTCATATGTTAGACAG	658
QY	661	TGAATGCGTATGAGATATGGAAGAACATCGATGGCGAAGCGACGCTCTTGCGTTAACACT	720
Db	659	TGAATGCGTATGAGATATGGAAGAACACCGATGGCGAAGCGACGCTCTTGCGTTAACACT	718
QY	721	GACCGTCAATGACAGAAAGCTGGGGAGCAAAACAGATTTAATATACCTGTATGTCACAGCC	780
Db	719	GACCGTCAATGACAGCAAAAGCTGGGGAGCAAAACAGATTTAATATACCTGTATGTCACAGCC	778
QY	781	CTAAACGATGTCATCTAGTTGTGTGGGCTTTATTTAGGCTGGTAAACGAAAGCTAACCGCGTGA	840
Db	779	CTAAACGATGTCATCTAGTTGTGTGGGCTTTATTTAGGCTGGTAAACGAAAGCTAACCGCGTGA	838
QY	841	AGTTGACCGCGCTGGGGAGTACGAGTCGCAAGATTAAAACTCAAGAAATTGACGGGGACCC	900
Db	839	AGTTGACCGCGCTGGGGAGTACGAGTCGCAAGATTAAAACTCAAGAAATTGACGGGGACCC	898
QY	901	GCACAAAGCGGTGATTAATGTGGAATTAATTTGATGCAACGCGAAAAACCTTAATCTTACCTT	960
Db	899	GCACAAAGCGGTGATTAATGTGGAATTAATTTGATGCAACGCGAAAAACCTTAATCTTACCTT	958
QY	961	GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTGGGGAACGCTAACAGGTTG	1017
Db	959	GACATGTAGCGAAGCCCGCGAGAGGTGATGTGCGGAAGGGAACGGCAACAGGTTG	1018
QY	1018	CTGATAGGCTGTGCGTCACTGATGTCGTGCGTAGATGTTGGTTAAGTCCCGCAACAGCGCA	1077
Db	1019	CTGATAGGCTGTGCGTCACTGATGTCGTGCGTAGATGTTGGTTAAGTCCCGCAACAGCGCA	1078
QY	1078	ACCTTGTTCATTAATTTGCAATCTTTGGTTGGGCACTTTAATGAGACTGCGGTTGACAA	1137
Db	1079	ACCTTGTTCATTAATTTGCAATCTTTGGGCACTTTAATGAAACGCGGTTGACAA	1138
QY	1138	CCGAGAGAAAGTGGGGAATGACGTCAAGTCTCATAGCCCTTAATGGGATGGGCTTCAACG	1197
Db	1139	CCGAGAGAAAGTGGGGAATGACGTCAAGTCTCATAGCCCTTAATGGGATGGGCTTCAACG	1198
QY	1198	TAAATCAATAGGCGGTACAGAGGTTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGC	1257
Db	1199	TAAATCAATAGGCGGTACAGAGGTTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGC	1258

QY	1258	GGGTGTAAGTCCGGATCCGAGTCTTGCACTGCACTCCGTGAAAGTCGAAATGCTAGTAAT	1317
DB	1259	GGCTGTATATCCGGATCCGAGTCTTGCACTGCACTCCGTGAAAGTCGAAATGCTAGTAAT <td>1318</td>	1318
QY	1318	CGCGGATAGCAATGTCGCGGGTGAATAGCTTCCGGGCTTTGTACACACCGCCGCTACAC	1377
DB	1319	CGCGGATAGCAATGTCGCGGGTGAATAGCTTCCGGGCTTTGTACACACCGCCGCTACAC	1378
QY	1378	CATGGAGTGGGATTTACACAGACGAGTAGTCTAACCTTAAGAGAGGGCGTTGCCACGG	1437
DB	1379	CATGGAGTGGGATTTACACAGACGAGTAGTCTAACCTTAAGAGAGGGCGCTGCCACGG	1438
QY	1438	TGAGATTCAATGACTGGGGTG	1457
DB	1439	TGAGATTCAATGACTGGGGTG	1458
RESULT 69			
LOCUS	AY123807	1498 bp	DNA linear BCT 23-SEP-2003
DEFINITION	Nitrosospirita multiformis 16S ribosomal RNA gene, partial sequence.		
ACCESSION	AY123807		
VERSION	AY123807.1	GI:24474436	
KEYWORDS			
SOURCE			
ORGANISM	Nitrosospirita multiformis ATCC 25196 (Nitrosolobus multiformis)		
	Nitrosospirita multiformis ATCC 25196		
	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;		
	Nitrosomonadaceae; Nitrosospirita.		
REFERENCE	1 (bases 1 to 1498)		
AUTHORS	Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roser,A. and Koops,H.P.		
TITLE	16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal of a new lineage within the nitrosomonads		
JOURNAL	Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)		
PUBMED	13130037		
REFERENCE	2 (bases 1 to 1498)		
AUTHORS	Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Rosser,A. and Wagner,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany		
FEATURES	Location/Qualifiers		
SOURCE	1..1498		
	/organism="Nitrosospirita multiformis ATCC 25196"		
	/mol_type="genomic DNA"		
	/strain="N13: C-71; ATCC 25196"		
	/db_xref="ATCC:25196"		
	/db_xref="Etaxon:323848"		
	/note="type strain"		
	synonym: Nitrosolobus multiformis"		
	<1..>1498		
	/product="16S ribosomal RNA"		
ORIGIN			
Query Match	87.8%;	Score 1279.8;	DB 1; Length 1498;
Best Local Similarity	93.6%;	Pred. No. 4.7e-13;	
Matches 1368;	Conservative 0;	Mismatches 87;	Indels 6; Gaps 3;
QY	1	ATTGAACGTGGCGGCAGTCCTTACATGCAATGCAAGTCGAACGGCAGCAGATGCTTGAT	60
DB	3	ATTGAACCTGGCGGCAGTCCTTACATGCAATGCAAGTCGAACGGCAGCAGATGCTTGAT	60
QY	61	CTGGTGGCGAGTGGCGGCAGCGGTGATGATGATGATGCAAGACGATCACAAGAGGGGGGTA	120
DB	61	CTGGTGGCGAGTGGCGGCAGCGGTGATGATGATGATGCAAGACGATCCTTAAGTGGGGGTA	120
QY	121	ACGCATCGAAGATGTGCTAATACCGCATATACCTTAAGAGAGAAAGCAGGGATCGAAA	180
DB	121	ACGCATCGAAGATGTGCTAATACCGCATATACCTCGAGAGAGAAAGCAGGGATCGCAA	180
QY	181	GACCTTGGCTTTGAGCGGCGCATGTCCTGATTAGCTAATTGGTGGGGTAAAGGCTTAC	240

Db	181	GACCTTGCGCTTTTGGAGCGGCCCATGTCTGATTAGTAAGTGGTAGATRAAGGCTTAC	240
Qy	241	CAAGCCGACGATCTAGTAGTGGTCTGAGAGGACGACGACCACTGAGACTGAGACG	300
Db	241	CAAGCCGTCGATCGATGATAGCTGTCTGAGAGGACGACGACCACTGAGACTGAGACG	300
Qy	301	GCCCAAGACTCTTAAGGAGAGGACGACGATGGGGAATTTTGGACAAATGGGCGCAAGCTGATC	360
Db	301	GCCCAAGACTCTTAAGGAGAGGACGACGATGGGGAATTTTGGACAAATGGGCGAAACCTGATC	360
Qy	361	CAGCAATGCGCGTAGAGTAAGAAAGGCGCTTGAGTTGTAAAGCTCTTTCAGTCGAGAGA	420
Db	361	CAGCCATGCGCGTAGAGTAAGAAAGGCGCTTGAGTTGTAAAGCTCTTTCAGCGCGAAGCA	420
Qy	421	AAAGGTTACGGTAAATATATCGTGACTAGACGGTATCGACAGAAAGACCGGCTTAC	480
Db	421	AAAGGTTACGGTAAATATATCGGTTATACCGTGACTAGCGTAACGGAAGAAAGCACCGGCTTAC	480
Qy	481	TACGTGCCAGACGCCGCGTAAATCGTAGGGTGCMAAGCTTAAATCGGAATTAATCTGGGCGT	540
Db	481	TACGTGCCAGACGCCGCGTAAATCGTAGGGTGTGACGCTTAAATCGGAATTAATCTGGGCGT	540
Qy	541	AAAGGGTCCGAGGCGGCG--CTTGTAGTCAGATGTGAAATCCCGGGCTTAACTGAGGA	599
Db	541	AAAGGGTCCGAGGCGGCGGCTTTTGTAGTCAGATGTGAAATCCCGGGCTTAACTGAGGA	600
Qy	600	TTGCGTTTGAATCTTACAAGGCTAGAGTGTGGCAGAGGAGGGTGGAAATTCATGTGTAGCA	659
Db	601	CTGGGTTTGAATCTTACAAGGCTAGAGTGTGGCAGAGGAGGGTGGAAATTCACGCTGTAGCA	660
Qy	660	GTGAATTCGTTAGAGATTTGGAAGAACATGCAATGGCCGAAAGGCAACCTCTGGGTTAACAC	719
Db	661	GTGAATTCGTTAGAGATGTGGAAGAACACCGATGGCGAAGGCAACCCCTGGGGTTAACAC	720
Qy	720	TGAGCGCTATGACGAAAGCGTGGGGAGCAAAAGATAGATACCTCGTAGTCCACGC	779
Db	721	TGAGCGCTATGAGGACGAAAGCGTGGGGAGCAAAAGATAGATACCTCGTAGTCCACGC	780
Qy	780	CCTAAACGATGTCACTAGATTGTGGGCTTATTAGGCTTGGTAAACGAAGCTTAAACGCGTG	839
Db	781	CCTAAACGATGTCAACTAGATTGTGGGCTTAAAGGACTTGGTAAACGACGCTTAAACGCGTG	840
Qy	840	AAAGTTGACCCGCTGGGGAGTACGGTCCGAAATTTAAATCTAAAGAAATTTGACGGGGAC	899
Db	841	AAAGTTGACCCGCTGGGGAGTACGGTCCGAAATTTAAATCTAAAGAAATTTGACGGGGAC	900
Qy	900	CGCAACAACGGGATTTATGTGATTTAATTTGATGCAACGGAAGAAACCTTACCTACCT	959
Db	901	CGCAACAACGGGATTTATGTGATTTAATTTGATGCAACGGAAGAAACCTTACCTACCT	960
Qy	960	TGACATGTAGCGAAATTTTCTAGAGATGATTAAGTGC--TTCCGGAGCGCTAACACAGGT	1016
Db	961	TGACATGTAGCGAGACCGCGCGAGAGGGGTGTGCGCGAAGGGAAACGCGAACACAGGT	1020
Qy	1017	GCTGATAGGCTGTGCTCAGCTTCGTGTGCTGATGATTTGGGTTAGTCCCGCAAGAGCG	1076
Db	1021	GCTGATAGGCTGTGCTCAGCTTCGTGTGCTGATGATTTGGGTTAGTCCCGCAAGAGCG	1080
Qy	1077	AAACCTTGTCAATTAATTTGSCATCAATTTGGTGGGCACTTTAATGATGACTCGGCTGACAA	1136
Db	1081	AAACCTTGTCAATTAATTTGSCATCAATTTGAGTTGGGCACTTTAATGATGACTCGGCTGACAA	1140
Qy	1137	ACCGAGAGAGTGGGAGTGAAGTCAAGTCTCTATGCGCTTATGGTAGGGCTTACAC	1196
Db	1141	ACCGAGAGAGTGGGAGTGAAGTCAAGTCTCTATGCGCTTATGGTAGGGCTTACAC	1200
Qy	1197	GTAATATCATGCGCGTACAGAGGTTTCCAAACCGCGAGGGGAGCTAATCTCAGAAAG	1256
Db	1201	GTAATATCATGCGCGTACAGAGGTTTCCAAACCGCGAGGGGAGCTAATCTCAGAAAG	1260
Qy	1257	CGCGTCGATGCCGAGTGGAGTCTGCAACTGCACTCCGTGAAGTCGGAATTCGCTAGTAA	1316
Db	1261	CGCGTCGATGCCGAGTGGAGTCTGCAACTGCACTCCGTGAAGTCGGAATTCGCTAGTAA	1320

QY 1317 TCGCGGATCAGCATGCTCGCGGTGAATACGTTCCCGGCTTGTATACACACCCCGCTCACA 1376  
DB 1321 TCGCGGATCAGCATGCTCGCGGTGAATACGTTCCCGGCTTGTATACACACCCCGCTCACA 1380  
QY 1377 CCATGGAGTGGGTTTCCACCAAGAGGTAGTCTTAACCGTAAAGAGGCGCGTGGCAGC 1436  
DB 1381 CCATGGAGTGGGTTTCCACCAAGAGGTAGTCTTAACCGCAAGAGGCGCGCTGGCAGC 1440  
QY 1437 GTGAGATTGATGCTGGGGTG 1457  
DB 1441 GTGAGATTGATGCTGGGGTG 1461

RESULT 70  
AY856376 1457 bp DNA linear BCT 15-JUN-2005  
LOCUS Nitrosospiira sp. PM2 16S ribosomal RNA gene, partial sequence.  
ACCESSION AY856376  
VERSION AY856376.1 GI:57338939  
KEYWORDS Nitrosospiira sp. PM2  
SOURCE Nitrosospiira sp. PM2  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
REFERENCE 1 (bases 1 to 1457)  
AUTHORS Takahashi, R., Yabe, R. and Tokuyama, T.  
TITLE Direct Submision  
JOURNAL Submitted (14-DEC-2004) Agricultural and Biological Chemistry,  
College of Bioresource Sciences, Nihon University, 1866 Kameino,  
Fujisawa, Kanagawa 252-8510, Japan

FEATURES  
source 1.1457  
/organism="Nitrosospiira sp. PM2"  
/mol\_type="genomic DNA"  
/strain="PM2"  
/db\_xref="taxon:308064"  
/note="similar to Nitrosospiira (Nitrosolobus) multiformis"  
<1..>1457  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 87.7%; Score 1277.2; DB 1; Length 1457;  
Best Local Similarity 93.6%; Pred. No. 5.3e-13;  
Matches 1366; Conservative 0; Mismatches 88; Indels 6; Gaps 3;

QY 1 ATTTGAACGCTGGGCGATGCTTTACATGCAAGTGAAGGAGGAGCGGATGCTTGCAT 60  
DB 1 ATTTGAACGCTGGGCGATGCTTTACATGCAAGTGAAGGAGGAGCGGATGCTTGCAT 58  
QY 61 CTGGTGGCGATGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 59 CTGGTGGCGATGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 118  
QY 121 ACGCATGAAAGATGTGCTAATAACCGCATATCTCTAAGAGAGAAACAGGAGATGAAA 180  
DB 119 ACGCATGAAAGATGTGCTAATAACCGCATATCTCTAAGAGAGAAACAGGAGATGCAA 178  
QY 181 GACCTTGGCGTTTGGAGCGGCGGATGCTGATTAAGTAACTTGTGGGTAAAGGCTTAC 240  
DB 179 GACCTTGGCGTTTGGAGCGGCGGATGCTGATTAAGTAACTTGTGGGTAAAGGCTTAC 238  
QY 241 CAAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 239 CAAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298  
QY 301 GCCCAGATCTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
DB 299 GCCCAGATCTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 358  
QY 361 CAGCATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
DB 359 CAGCATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418

QY 421 AAAGTTACGGTAAATATCGNACTCATGACGGTATCGACAGAGAACACCGGCTAAC 480  
DB 419 AACGGTACGGTAAATATCGNACTCATGACGGTATCGACAGAGAACACCGGCTAAC 478  
QY 481 TACGTGCCAGACCGCGGATTAATAGAGGTGACAGGCTTAATCGAATTAAGTGGCGT 540  
DB 479 TACGTGCCAGACCGCGGATTAATAGAGGTGACAGGCTTAATCGAATTAAGTGGCGT 538  
QY 541 AAAGGTGCGCAGGCGGCTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 539 AAAGGTGCGCAGGCGGCTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 598  
QY 601 TGGCTTTGAAATCAAGAGCTAGATGTGCGAGAGAGGTGATTAATTCATGTTAGCAG 660  
DB 599 TGGCTTTGAAATCAAGAGCTAGATGTGCGAGAGAGGTGATTAATTCATGTTAGCAG 658  
QY 661 TGAATGCGTAGAGATATGAGAGAACATCGATGCGAAGGAGCGCTCTGGGTTAACT 720  
DB 659 TGAATGCGTAGAGATATGAGAGAACATCGATGCGAAGGAGCGC-CTGGGTTAACT 717  
QY 721 GACGCTATGACGAAAGCGTGGGAGCAAAAGGATTAAGATACCTGCTAGTCCAGCGC 780  
DB 718 GACGCTATGACGAAAGCGTGGGAGCAAAAGGATTAAGATACCTGCTAGTCCAGCGC 777  
QY 781 CTAAACGATGCACTAGTGTGGGCTTATTAAGGCTTGTATGCAAGCTAACGCGTGA 840  
DB 778 CTAAACGATGCACTAGTGTGGGCTTATTAAGGCTTGTATGCAAGCTAACGCGTGA 837  
QY 841 AGTTGACGCGCTGGGAGATACCGTGCAGAGTTAAACTCAAGAGATTGACGGGAGCC 900  
DB 838 AGTTGACGCGCTGGGAGATACCGTGCAGAGTTAAACTCAAGAGATTGACGGGAGCC 897  
QY 901 GCACAAAGCGTGGATTAATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 960  
DB 898 GCACAAAGCGTGGATTAATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 957  
QY 961 GACATGTACGAAATTTTCTAGATATGATTAATGATGATGATGATGATGATGATGATGAT 1017  
DB 958 GACATGTACGAAATTTTCTAGATATGATTAATGATGATGATGATGATGATGATGATGAT 1017  
QY 1018 CTGCATGCTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077  
DB 1018 CTGCATGCTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077  
QY 1078 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
DB 1078 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
QY 1138 CCGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
DB 1138 CCGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
QY 1198 TAAATCAATGAGCGCTGACAGAGGCTTCCAAACCGGAGGAGGAGCTTAATTCAGAAAG 1257  
DB 1198 TAAATCAATGAGCGCTGACAGAGGCTTCCAAACCGGAGGAGGAGCTTAATTCAGAAAG 1257  
QY 1258 GCGTGTAGTCCGATCGAGTCTGCACTGCTCGTGAAGTGGATGCTGATGAT 1317  
DB 1258 GCGTGTAGTCCGATCGAGTCTGCACTGCTCGTGAAGTGGATGCTGATGAT 1317  
QY 1318 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377  
DB 1318 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377  
QY 1378 CATGGAGAGGAGTTTCAACGAGAGAGTATGTTAAACCGAGAGAGGCGCTGCCACG 1437  
DB 1378 CATGGAGAGGAGTTTCAACGAGAGAGTATGTTAAACCGAGAGAGGCGCTGCCACG 1437  
QY 1438 TGAGATTCAATGACTGGGGTG 1457  
DB 1438 TGAGATTCAATGACTGGGGTG 1457

Search completed: May 6, 2006, 19:45:17  
Job time : 7523 secs

---







## ORIGIN

Query Match 99.8%; Score 1453.8; DB 3; Length 1457;  
Best Local Similarity 99.9%; Pred. No. 9.2e-16;  
Matches 1455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATTGAACGCTGCGCGGACCTTTTACATGCAAGTCGACGCGACGCGATCTTGAT 60
DB 1 ATTGAACGCTGCGCGGACCTTTTACATGCAAGTCGACGCGACGCGATCTTGAT 60
QY 61 CTGTGGCCGAGTGGCGGACGCGGTGAGTAATGCAATCGGAAGTAAGAGGGGGTA 120
DB 61 CTGTGGCCGAGTGGCGGACGCGGTGAGTAATGCAATCGGAAGTAAGAGGGGGTA 120
QY 121 AGGCATCGAAAGATGTCTAATACCGCATATACCTTAAGAGGAAAGCAGGGATCGAA 180
DB 121 AGGCATCGAAAGATGTCTAATACCGCATATACCTTAAGAGGAAAGCAGGGATCGAA 180
QY 181 GACCTTGCGCTTTTGGAGCGGCGGATGCTGATTAGCTAGTGGTGGGTAAAGCCTTAC 240
DB 181 GACCTTGCGCTTTTGGAGCGGCGGATGCTGATTAGCTAGTGGTGGGTAAAGCCTTAC 240
QY 241 CAAGGCGACATCATGTAAGTGTGTTGAGAGAGCAACGCACTGGGACTGAGACG 300
DB 241 CAAGGCGACATCATGTAAGTGTGTTGAGAGAGCAACGCACTGGGACTGAGACG 300
QY 301 GCCCAGACTCTTAACGCGGAGCAGAGTGGGAAATTTTGAACAATGGGCGCAGCTGATC 360
DB 301 GCCCAGACTCTTAACGCGGAGCAGAGTGGGAAATTTTGAACAATGGGCGCAGCTGATC 360
QY 361 CAGCAATGCGCGCTGAGTGAAGAAAGCCTTCGGGTGTAAGCTCTTTCAGTCGAGAA 420
DB 361 CAGCAATGCGCGCTGAGTGAAGAAAGCCTTCGGGTGTAAGCTCTTTCAGTCGAGAA 420
QY 421 AAAGCTTAACGCTTAATATGCTGACCAATGACGCTATCGACAGAAAGCAACGCGCTTAC 480
DB 421 AAAGCTTAACGCTTAATATGCTGACCAATGACGCTATCGACAGAAAGCAACGCGCTTAC 480
QY 481 TAGCTGCGACGAGCGCGGTAAATCGTAGGCTGCAACGCTTAATCGAATTTACTGGCGCT 540
DB 481 TAGCTGCGACGAGCGCGGTAAATCGTAGGCTGCAACGCTTAATCGAATTTACTGGCGCT 540
QY 541 AAAGGTGCGCAGCGCGCTTTGTAAGTCAGATGTAATCCCGGGCTTAACTTGGGAAT 600
DB 541 AAAGGTGCGCAGCGCGCTTTGTAAGTCAGATGTAATCCCGGGCTTAACTTGGGAAT 600
QY 601 TGCCTTTGAACTCAAAAGCTAGAGTGGCAGAGGGAGTGGAATTCATGTGTAGCAG 660
DB 601 TGCCTTTGAACTCAAAAGCTAGAGTGGCAGAGGGAGTGGAATTCATGTGTAGCAG 660
QY 661 TGAATATCGCTAGATATGGAAGAAATCGATGGCGAAAGGAGCCTCTGCGGTTAACT 720
DB 661 TGAATATCGCTAGATATGGAAGAAATCGATGGCGAAAGGAGCCTCTCTGCGGTTAACT 720
QY 721 GACGCTCATGCAAAAAGCTGCGGAGCAAAAGATTAAGTACCTTGTAGTCCAGCC 780
DB 721 GACGCTCATGCAAAAAGCTGCGGAGCAAAAGATTAAGTACCTTGTAGTCCAGCC 780
QY 781 CTAACGATGTCATCTAGTTGTTGGCCTTATTAGGCTTGGTAAAGAGCTTAACGCGTGA 840
DB 781 CTAACGATGTCATCTAGTTGTTGGCCTTATTAGGCTTGGTAAAGAGCTTAACGCGTGA 840
QY 841 AGTTGACCGCGCTGGGAGATGAGGTGCAAGATTAATACTCAAGGAATTGACGGGAGCC 900
DB 841 AGTTGACCGCGCTGGGAGATGAGGTGCAAGATTAATACTCAAGGAATTGACGGGAGCC 900
QY 901 GCACAAAGCGGTGATTAATGATTAATTCATGCAACGCGAAAACTTACCTT 960
DB 901 GCACAAAGCGGTGATTAATGATTAATTCATGCAACGCGAAAACTTACCTT 960
QY 961 GACATGTAGCGAATTTTCTAGAGATAGTATGCTTGGGAAAGCTTAACAGGTGTG 1020
DB 961 GACATGTAGCGAATTTTCTAGAGATAGTATGCTTGGGAAAGCTTAACAGGTGTG 1020
```

```
QY 1021 CATGCTGTCTGACAGTCTGTGCTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACC 1080
DB 1021 CATGCTGTCTGACAGTCTGTGCTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACC 1080
QY 1081 CTGTGATTAATTTGCCATCATTTGTTGGGCACTTTAATAGACCTGCGGTGCAAAACG 1140
DB 1081 CTGTGATTAATTTGCCATCATTTGTTGGGCACTTTAATAGACCTGCGGTGCAAAACG 1140
QY 1141 GAGGAAGTGGGGATGACGTCAAGTCTCATGCGCTTATGGTGGGCTTCAACAGTAA 1200
DB 1141 GAGGAAGTGGGGATGACGTCAAGTCTCATGCGCTTATGGTGGGCTTCAACAGTAA 1200
QY 1201 TACAAATGCGCGCTACAGAGGTTGCAACCCCGAGGGGAGCTAATCTCAGAAAGCGG 1260
DB 1201 TACAAATGCGCGCTACAGAGGTTGCAACCCCGAGGGGAGCTAATCTCAGAAAGCGG 1260
QY 1261 TCGTAGTCCGGAATCGAGTCTGCACTCGACTCCGTAAGTCGGAATCGCTAATTCGC 1320
DB 1261 TCGTAGTCCGGAATCGAGTCTGCACTCGACTCCGTAAGTCGGAATCGCTAATTCGC 1320
QY 1321 GGATCAGATGTCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCCTGACACAT 1380
DB 1321 GGATCAGATGTCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCCTGACACAT 1380
QY 1381 GGGAGTGGGTTTACCAAGAGCGTAACTTAACCGTAAGAGAGGGCGCTTGCCACGGTGA 1440
DB 1381 GGGAGTGGGTTTACCAAGAGCGTAACTTAACCGTAAGAGAGGGCGCTTGCCACGGTGA 1440
QY 1441 GATTCAATGACTGGGGTG 1457
DB 1441 GATTCAATGACTGGGGTG 1457
```

RESULT 3  
AF386753 1457 bp DNA linear ENV 10-MAY-2004  
LOCUS AF386753  
DEFINITION Nitrosomonas sp. R7c140 16S ribosomal RNA gene, partial sequence.  
ACCESSION AF386753  
VERSION AF386753.1 GI:17864830  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. R7c140  
ORGANISM Nitrosomonas sp. R7c140  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas; environmental samples.

REFERENCE  
AUTHORS 1 (bases 1 to 1457)  
Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)

JOURNAL PUBMED 11722936  
REFERENCE 2 (bases 1 to 1457)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condon Dr., Moorpark, CA 93021, USA  
LOCATION/Qualifiers

## FEATURES

source  
1. 1457  
/organism="Nitrosomonas sp. R7c140"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichments"  
/db\_xref="taxon:180543"  
/clone="R7c140"  
/environmental\_sample  
<1..>1457  
/product="16S ribosomal RNA"

## ORIGIN

Query Match 99.7%; Score 1452.2; DB 3; Length 1457;  
Best Local Similarity 99.8%; Pred. No. 9.7e-16;  
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATTGAACGCTGGCGGATGCTTTACATGCAAGTCGAAACGCGAGCACGGATCTTGCAAT 60  
Db 1 ATTGAACGCTGGCGGATGCTTTACATGCAAGTCGAAACGCGAGCACGGATCTTGCAAT 60  
QY 61 CTGGTGGCGAGTGGCGGACCGGATGAGTAAATGCAATCGGAACGTAATCCAGAAAGAGGGGGGTA 120  
Db 61 CTGGTGGCGAGTGGCGGACCGGATGAGTAAATGCAATCGGAACGTAATCCAGAAAGAGGGGGGTA 120  
QY 121 ACGCATCGAAAGATGTGCTTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGGATCGAAA 180  
Db 121 ACGCATCGAAAGATGTGCTTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGGATCGAAA 180  
QY 181 GACCTTGCGCTTTTGGAGCGCGCATGTCTGATTAAGTAGTGTGGGGGTAAAGGCGCTAC 240  
Db 181 GACCTTGCGCTTTTGGAGCGCGCATGTCTGATTAAGTAGTGTGGGGGTAAAGGCGCTAC 240  
QY 241 CAAGGCGACGATCAGTAAATGTTGCTGAGAGAGACGACCGACACTGGGACTGAGACAG 300  
Db 241 CAAGGCGACGATCAGTAAATGTTGCTGAGAGAGACGACCGACACTGGGACTGAGACAG 300  
QY 301 GCCCAACTCTTAACGGGAGGACAGCAATGGGGAAATTTTGGACATATGGGCGCAAGCTGATC 360  
Db 301 GCCCAACTCTTAACGGGAGGACAGCAATGGGGAAATTTTGGACATATGGGCGCAAGCTGATC 360  
QY 361 CAGCAATGCGCGCTGAGTGAAGAAAGGCTTCGGGTTGTAAAGCTCTTCAGTGAAGAA 420  
Db 361 CAGCAATGCGCGCTGAGTGAAGAAAGGCTTCGGGTTGTAAAGCTCTTCAGTGAAGAA 420  
QY 421 AAAGGTTACGTTAAATATCTGTGACCCATGACGGTAATCGACAGAAAGACACCGGCTAAC 480  
Db 421 AAAGGTTACGTTAAATATCTGTGACCCATGACGGTAATCGACAGAAAGACACCGGCTAAC 480  
QY 481 TACGTGCCACAGCCCGCGTAAATACGTAAAGGTCACAGCGTTAATCTGGAATTTCTGGGCGT 540  
Db 481 TACGTGCCACAGCCCGCGTAAATACGTAAAGGTCACAGCGTTAATCTGGAATTTCTGGGCGT 540  
QY 541 AAAGGGTGGCAGCGCGCTTGTAACTAGTGAATGTCGAAATCCCGGGGCTTAACCTGGGAAT 600  
Db 541 AAAGGGTGGCAGCGCGCTTGTAACTAGTGAATGTCGAAATCCCGGGGCTTAACCTGGGAAT 600  
QY 601 TSCGTTTGAACCTAACAGCTAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660  
Db 601 TSCGTTTGAACCTAACAGCTAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660  
QY 661 TGAATATGCTAGATATGAGAGAAACATGATGGCGAAGCGACCTCTCGGGTTAACT 720  
Db 661 TGAATATGCTAGATATGAGAGAAACATGATGGCGAAGCGACCTCTCGGGTTAACT 720  
QY 721 GACGCTCATGCAAGAAAGGTGGGAGCAAAACAGATTAAGTATCCCTGGTAAGTCCAGCGC 780  
Db 721 GACGCTCATGCAAGAAAGGTGGGAGCAAAACAGATTAAGTATCCCTGGTAAGTCCAGCGC 780  
QY 781 CTAAACGATGTCAACTAGTGTGGGCTTAATTAAGGCTTGTATGACAAAGCTAAACGGGTGA 840  
Db 781 CTAAACGATGTCAACTAGTGTGGGCTTAATTAAGGCTTGTATGACAAAGCTAAACGGGTGA 840  
QY 841 AGTTAGACCGCTCGGGAGATACGCTGCAAGATTTAAACCTCAAGAGAAATGACGGGGACCC 900  
Db 841 AGTTAGACCGCTCGGGAGATACGCTGCAAGATTTAAACCTCAAGAGAAATGACGGGGACCC 900  
QY 901 GCACAGACGGTGAATTAATGATTAATTCGATGCAACGCGAAGAACTTACCTACCTT 960  
Db 901 GCACAGACGGTGAATTAATGATTAATTCGATGCAACGCGAAGAACTTACCTACCTT 960  
QY 961 GACATGTAGCGAATTTTCTAGAGATTAAGTAAATGCTTCGGGAAACGCTAAACAGAGTGTCTG 1020  
Db 961 GACATGTAGCGAATTTTCTAGAGATTAAGTAAATGCTTCGGGAAACGCTAAACAGAGTGTCTG 1020  
QY 1021 CATGGCTGTCTGACGCTGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1080  
Db 1021 CATGGCTGTCTGACGCTGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1080  
QY 1081 CTTGTCAATTAATGCGATCAATTTGGTGGGCACTTTAATGAGACATGCGCGGTACAAACCG 1140

Db 1081 CTTGTCAATTAATGCGATCAATTTGGTGGGCACTTTAATGAGACTGCCGTACAAACCG 1140  
QY 1141 GAGGAAGTGGCGATGACCTCAAGTCTCTATGCGCTTATGGGTAGGGCTTACACGTAA 1200  
Db 1141 GAGGAAGTGGCGATGACCTCAAGTCTCTATGCGCTTATGGGTAGGGCTTACACGTAA 1200  
QY 1201 TACAAATGGCGGTACAGAGGGGTGGCAACCGCGAGGGGGAGCTAATCTCAGAAAGCGCG 1260  
Db 1201 TACAAATGGCGGTACAGAGGGGTGGCAACCGCGAGGGGGAGCTAATCTCAGAAAGCGCG 1260  
QY 1261 TCGTAGTCGGATTCGAGAGTCTGCAACTGCACTCCGTGAAGTCTGAATGCTAATGCTC 1320  
Db 1261 TCGTAGTCGGATTCGAGAGTCTGCAACTGCACTCCGTGAAGTCTGAATGCTAATGCTC 1320  
QY 1321 GGATCAGCATGTGCGGTGATATGCTTCCGGGCTTTGTACACACCGCCCGTACACCAT 1380  
Db 1321 GGATCAGCATGTGCGGTGATATGCTTCCGGGCTTTGTACACACCGCCCGTACACCAT 1380  
QY 1381 GGGAGTGGGTTTACACAGAGAGGTAGTCTAACGTAAGAGAGGGGCGCTTGCACAGTGA 1440  
Db 1381 GGGAGTGGGTTTACACAGAGAGGTAGTCTAACGTAAGAGAGGGGCGCTTGCACAGTGA 1440  
QY 1441 GATTCATGACTGGGGTG 1457  
Db 1441 GATTCATGACTGGGGTG 1457  
  
RESULT 4  
CQ796908 1457 bp DNA linear PAT 19-APR-2004  
LOCUS CQ796908  
DEFINITION Sequence 1 from Patent WO2004026772.  
ACCESSION CQ796908  
VERSION CQ796908.1 GI:46408534  
KEYWORDS  
SOURCE  
ORGANISM  
unidentified  
unclassified sequences.  
REFERENCE  
1 Hovanec, T. A.  
Ammonia-oxidizing bacteria and methods of using and detecting  
the same  
Patent: WO 2004026772-A 1 01-APR-2004;  
JOURNAL  
Aqualia Inc. (US)  
FEATURES  
source  
Location/Qualifiers  
1..1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="NOB Type A R7clone140 16S rDNA"  
  
ORIGIN  
Query Match 99.7%; Score 1452.2; DB 6; Length 1457;  
Best Local Similarity 99.8%; Pred. No. 9.7e-16;  
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 ATTGAACGCTGGCGGATGCTTTACATGCAAGTCGAAACGCGAGCACGGATCTTGCAAT 60  
Db 1 ATTGAACGCTGGCGGATGCTTTACATGCAAGTCGAAACGCGAGCACGGATCTTGCAAT 60  
QY 61 CTGGTGGCGAGTGGCGGACCGGATGAGTAAATGCAATCGGAACGTAATCCAGAAAGAGGGGGGTA 120  
Db 61 CTGGTGGCGAGTGGCGGACCGGATGAGTAAATGCAATCGGAACGTAATCCAGAAAGAGGGGGGTA 120  
QY 121 ACGCATCGAAAGATGTGCTTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGGATCGAAA 180  
Db 121 ACGCATCGAAAGATGTGCTTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGGATCGAAA 180  
QY 181 GACCTTGCGCTTTTGGAGCGCGCATGTCTGATTAAGTAGTGTGGGGGTAAAGGCGCTAC 240  
Db 181 GACCTTGCGCTTTTGGAGCGCGCATGTCTGATTAAGTAGTGTGGGGGTAAAGGCGCTAC 240  
QY 241 CAAGGCGACGATCAGTAAATGTTGCTGAGAGAGACGACCGACACTGGGACTGAGACAG 300

Db	241	CAAGCGACGATCAGTAGTGGTCTTGAGAGACACACGACACTGAGACTAGACACG	300
Qy	301	GCCGACACTCTTAAGGAGGAGGACGAGTGGGGAATTTTGGACAAATGGGCGCAACCTGATC	360
Db	301	GCCGACACTCTTAAGGAGGAGGACGAGTGGGGAATTTTGGACAAATGGGCGCAACCTGATC	360
Qy	361	CAGCAATGCCGCGTGAAGGAAGAGGCCCTTCGGGTGTAAAGCTCTTTCACTGAGAGAG	420
Db	361	CAGCAATGCCGCGTGAAGGAAGAGGCCCTTCGGGTGTAAAGCTCTTTCACTGAGAGAG	420
Qy	421	AAAGGTTACGGTAAATTAATCGTGAACCTATGACGAGTATCGAACGAAGAGACCGGGCTAAC	480
Db	421	AAAGGTTACGGTAAATTAATCGTGAACCTATGACGAGTATCGAACGAAGAGACCGGGCTAAC	480
Qy	481	TACGTCGACGAGCGCGCGGTAAATACGTAGAGGTGCAACGCTTAATCGAATTTACTGGCGGT	540
Db	481	TACGTCGACGAGCGCGCGGTAAATACGTAGAGGTGCAACGCTTAATCGAATTTACTGGCGGT	540
Qy	541	AAAGGTCGCGAGGCGGGCTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTGGGAAT	600
Db	541	AAAGGTCGCGAGGCGGGCTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTGGGAAT	600
Qy	601	TGCGTTTGAATCTAACAAAGCTTAGAGTGTGGAGAGGAGGTGAATTCAGATGTAGACAG	660
Db	601	TGCGTTTGAATCTAACAAAGCTTAGAGTGTGGAGAGGAGGTGAATTCAGATGTAGACAG	660
Qy	661	TGAATGCGTAGAGATATGAAGAACATCGATGCGCAAGGACGCTCTGGGTATACACT	720
Db	661	TGAATGCGTAGAGATATGAAGAACATCGATGCGCAAGGACGCTCTGGGTATACACT	720
Qy	721	GACGCTCAGCAGCAAGAACGCTGGGAGGCAACAGATTAAGTATACCTTGATGTCACGCC	780
Db	721	GACGCTCAGCAGCAAGAACGCTGGGAGGCAACAGATTAAGTATACCTTGATGTCACGCC	780
Qy	781	CTAAACGATGTCACCTAAGTTGTGGGCTTATTAGGCTTGSTTACGAAGCTTAACGCGTGA	840
Db	781	CTAAACGATGTCACCTAAGTTGTGGGCTTATTAGGCTTGSTTACGAAGCTTAACGCGTGA	840
Qy	841	AGTTAACCGCCCTGGGAGTACGCTGGCAAGATTAAACTCAAGAAATTGACGGGAGCC	900
Db	841	AGTTAACCGCCCTGGGAGTACGCTGGCAAGATTAAACTCAAGAAATTGACGGGAGCC	900
Qy	901	GCACAGCGGTGAGATTATGTGAGTTAATTCGATGACGACGCGAATAACTTACCTT	960
Db	901	GCACAGCGGTGAGATTATGTGAGTTAATTCGATGACGACGCGAATAACTTACCTT	960
Qy	961	GACATGTAGCAATTTTCTAAGAGATAGATTAGTCTTGGGAGCGCTAACACAGTCTGT	1020
Db	961	GACATGTAGCAATTTTCTAAGAGATAGATTAGTCTTGGGAGCGCTAACACAGTCTGT	1020
Qy	1021	CATGGCTGTGTCAGCTGCTGTGCTGAGATGTGGGTAAAGTCCCGCAACGAGCGCAAC	1080
Db	1021	CATGGCTGTGTCAGCTGCTGTGCTGAGATGTGGGTAAAGTCCCGCAACGAGCGCAAC	1080
Qy	1081	CTTGTCAATTAATTCGCATCATTTGGTGGGCACTTTAATGAGACTGCGGTGACAAACCG	1140
Db	1081	CTTGTCAATTAATTCGCATCATTTGGTGGGCACTTTAATGAGACTGCGGTGACAAACCG	1140
Qy	1141	GAGGAAGGTGGGAGTGAAGTCAAGTCTCATATGGCCCTTAATGGGTAGGGCTTACACGTTAA	1200
Db	1141	GAGGAAGGTGGGAGTGAAGTCAAGTCTCATATGGCCCTTAATGGGTAGGGCTTACACGTTAA	1200
Qy	1201	TACAAATGGGCGGTACAGAGGTGTCCAACCCGCGAGGGGAGACTAATCTCAGAAAGCGCG	1260
Db	1201	TACAAATGGGCGGTACAGAGGTGTCCAACCCGCGAGGGGAGACTAATCTCAGAAAGCGCG	1260
Qy	1261	TGCTAGTCCGAGTCCGAGTCTGCAACTCGACTCCGTAAGTGGAGTGGTAGTATGTC	1320
Db	1261	TGCTAGTCCGAGTCCGAGTCTGCAACTCGACTCCGTAAGTGGAGTGGTAGTATGTC	1320
Qy	1321	GGATGACATGTGCGGTGAATACGTTCCCGGCTTGTAGCACACGCCCGGTACACAT	1380
Db	1321	GGATGACATGTGCGGTGAATACGTTCCCGGCTTGTAGCACACGCCCGGTACACAT	1380

Db	1321	GGATCAGCATGTCCGGGTGAATACGTTCCCGGGCTTTGTACACACCGCCGTACACCAT	1387
Qy	1381	GGGAGTGGGGTTTCCACGAAGCAGTAGTCTAACCGTAAAGAGGGCCCTTGCCACGGTGA	1440
Db	1381	GGGAGTGGGGTTTCCACGAAGCAGTAGTCTAACCGTAAAGAGGGCCCTTGCCACGGTGA	1440
Qy	1441	GATTTCATGACTGGGGTG	1457
Db	1441	GATTTCATGACTGGGGTG	1457
RESULT 5			
CS089154			
LOCUS	CS089154	1457 bp	DNA linear PAT 25-MAY-2005
DEFINITION	Sequence 1 from Patent EP1502948.		
ACCESSION	CS089154		
VERSION	CS089154.1	GI:66714438	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE	1	unidentified unidentified unclassified.	
AUTHORS	Hovaneec, T. A. and Burrell, P. C.		
TITLE	Ammonia-oxidizing bacteria		
JOURNAL	Patent: EP 1502948-A 1 02-FEB-2005;		
FEATURES	Aquaria Inc. (US)		
source	Location/Qualifiers		
	1. 1457		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
	/note="Description of Unknown Organism: AOB Type A ammonia-oxidizing bacterium represented by K7 clone140"		
ORIGIN			
Query Match	99.7%;	Score 1452.2;	DB 6; Length 1457;
Best Local Similarity	99.8%;	Pred. No. 9.7e-16;	
Matches 1454; Conservative	0;	Mismatches 3;	Indels 0; Gaps 0
Qy	1	ATTGAACCTGGCGGATGCTTTACATGCAATGCAAGTCGAAACGGACGACGATGCTTGAT	60
Db	1	ATTGAACCTGGCGGATGCTTTACATGCAATGCAAGTCGAAACGGACGACGATGCTTGAT	60
Qy	61	CTGTGGCGGAGTGGCGGACGGGTAGTAAATGATCGTGGAAACGTATCCAAAGAGGGGGGT	120
Db	61	CTGTGGCGGAGTGGCGGACGGGTAGTAAATGATCGTGGAAACGTATCCAAAGAGGGGGGT	120
Qy	121	ACGCATCGAAGATGTCCTAATACCGCATATCTCTAAGAGGAAGACGGGGATCGAA	180
Db	121	ACGCATCGAAGATGTCCTAATACCGCATATCTCTAAGAGGAAGACGGGGATCGAA	180
Qy	181	GACCTTGGCGCTTTGGAAGCGGCGATGTCGTAGTAACTGATTGGTGGGGTAAAGGCTTAC	240
Db	181	GACCTTGGCGCTTTGGAAGCGGCGATGTCGTAGTAACTGATTGGTGGGGTAAAGGCTTAC	240
Qy	241	CAAGCGACATCACTATTGTTCTCGAAGAGACGACACGACCACTGGGGATCGAGACG	300
Db	241	CAAGCGACATCACTATTGTTCTCGAAGAGACGACACGACCACTGGGGATCGAGACG	300
Qy	301	GCCCAAGATCTCTAAGCGGAGGACGAGTGGGGAATTTGGACATGGGCGCAAGCCTGATC	360
Db	301	GCCCAAGATCTCTAAGCGGAGGACGAGTGGGGAATTTGGACATGGGCGCAAGCCTGATC	360
Qy	361	CAGCAATGCGCGGTAGTGAAGAGAGGCTTTGGGTTGTAAAGCTCTTTTCAATCGAGAA	420
Db	361	CAGCAATGCGCGGTAGTGAAGAGAGGCTTTGGGTTGTAAAGCTCTTTTCAATCGAGAA	420
Qy	421	AAAGTTTACGTTAAATCCGTGACCCATGACGATGACGAGAAAGACCCGGCTAAC	480
Db	421	AAAGTTTACGTTAAATCCGTGACCCATGACGATGACGAGAAAGACCCGGCTAAC	480
Qy	481	TACGTCCAGACAGCGCGGTAAATAGTAGGGTGAAGCGTTAATCGAATTAATCTGGGCGT	540
Db	481	TACGTCCAGACAGCGCGGTAAATAGTAGGGTGAAGCGTTAATCGAATTAATCTGGGCGT	540

```
Db      481  TAGGTGCACGACCGCGGTAAATACGTAGAGGTGCAAGCTTTAATCGAAATTAAGTGGGCGT 540
Qy      541  AAAAGGTGGCAGGCGGCTTTGTAATCAGATGTGAATTCCTCGGGCTTAAGTGGGAT 600
Db      541  AAAAGGTGGCAGGCGGCTTTGTAATCAGATGTGAATTCCTCGGGCTTAAGTGGGAT 600
Qy      601  TGGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660
Db      601  TGGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660
Qy      661  TGAATGCGTAGAGATATGAAAGACATCGATGCGCAAGCGACCTCTGGGTTAACT 720
Db      661  TGAATGCGTAGAGATATGAAAGACATCGATGCGCAAGCGACCTCTGGGTTAACT 720
Qy      721  GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATTCCTGGTATGTCACGCC 780
Db      721  GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATTCCTGGTATGTCACGCC 780
Qy      781  CTAAACGATGTCAACTAGTTGTTGGGCTTTATTAAGGCTTGATACGAGCTTAACGCGTGA 840
Db      781  CTAAACGATGTCAACTAGTTGTTGGGCTTTATTAAGGCTTGATACGAGCTTAACGCGTGA 840
Qy      841  AGTTGACCGCTGGGGAGTACGGTCCGAAGATTAAACTCAAGGAATTGACGGGGACCC 900
Db      841  AGTTGACCGCTGGGGAGTACGGTCCGAAGATTAAACTCAAGGAATTGACGGGGACCC 900
Qy      901  GCACAAAGCGGTGATTTATGTGATTTAATTCATGCAACCGCAAAAACCTTAACCTT 960
Db      901  GCACAAAGCGGTGATTTATGTGATTTAATTCATGCAACCGCAAAAACCTTAACCTT 960
Qy      961  GACATGTAGCAATTTTCTAGAGATAGATTAGTCTTCGGGAACGCTAACACAGGTGCTG 1020
Db      961  GACATGTAGCAATTTTCTAGAGATAGATTAGTCTTCGGGAACGCTAACACAGGTGCTG 1020
Qy      1021  CATGCTGTCTGACGCTGTCTGTGAGATGTTGGTTAAGTCCGCAACGAGCGCAACC 1080
Db      1021  CATGCTGTCTGACGCTGTCTGTGAGATGTTGGTTAAGTCCGCAACGAGCGCAACC 1080
Qy      1081  CTTGTCAATTAAATTTGTCATCTTTGTTGGGCACTTTAATGAGACTGCGCGTGAACAAACCG 1140
Db      1081  CTTGTCAATTAAATTTGTCATCTTTGTTGGGCACTTTAATGAGACTGCGCGTGAACAAACCG 1140
Qy      1141  GAGGAAGGTGGGATGATCAAGTCTCATGCGCTTTATGGGTAGGGCTTCAACGTAA 1200
Db      1141  GAGGAAGGTGGGATGATCAAGTCTCATGCGCTTTATGGGTAGGGCTTCAACGTAA 1200
Qy      1201  TACAATGGCGCGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAACGCG 1260
Db      1201  TACAATGGCGCGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAACGCG 1260
Qy      1261  TGTGTAGTCCGGATCCGAGTCTGCAACTCGACTCCGTAAGTGGGAATTCCTAGTATCCG 1320
Db      1261  TGTGTAGTCCGGATCCGAGTCTGCAACTCGACTCCGTAAGTGGGAATTCCTAGTATCCG 1320
Qy      1321  GGATCAGCATGTGCGGGTGAATACGTTCCGGGCTTTGTACACACCGCCCGTCAACCAT 1380
Db      1321  GGATCAGCATGTGCGGGTGAATACGTTCCGGGCTTTGTACACACCGCCCGTCAACCAT 1380
Qy      1381  GGGAGTGGGTTTACCAAGAGCAGGTAGTCTAAACCGTAAAGAGGGCGCTTGCACGCTGA 1440
Db      1381  GGGAGTGGGTTTACCAAGAGCAGGTAGTCTAAACCGTAAAGAGGGCGCTTGCACGCTGA 1440
Qy      1441  GATTTCATGACTGGGGTG 1457
Db      1441  GATTTCATGACTGGGGTG 1457

RESULT 5
CS089155      1457 bp      DNA      linear      PAT 25-MAY-2005
LOCUS      CS089155      Sequence 2 from Patent EP1502948.
DEFINITION      CS089155
ACCESSION      CS089155.1      GI:66714439
```

```
KEYWORDS      .
SOURCE      unidentified
ORGANISM      unidentified
REFERENCE      1
AUTHORS      Hovanec,T.A. and Burrell,P.C.
TITLE      Ammonia-oxidizing bacteria
JOURNAL      Patent: EP 1502948-A 2 02-FEB-2005;
              Aquaria Inc. (US)
FEATURES
  source      Location/Qualifiers
              1..1457
              /organism="unidentified"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32644"
              /note="Description of Unknown Organism: AOB Type A1
              ammonia-oxidizing bacterium represented by R7 clone187"
ORIGIN
  Query Match      99.7%; Score 1452.2; DB 6; Length 1457;
  Best Local Similarity 99.8%; Pred. No. 9,7e-16;
  Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1  ATTGAACGCTGGCGGCATGCTTTACATGCAATGCAAGTCGAACGCGACGACGATGCTTCAT 60
Db      1  ATTGAACGCTGGCGGCATGCTTTACATGCAATGCAAGTCGAACGCGACGACGATGCTTCAT 60
Qy      61  CTGGTGGGAGTGGCGGACGGGTGATATGATCGGAACGTAATCGAAGAGGGGGTA 120
Db      61  CTGGTGGGAGTGGCGGACGGGTGATATGATCGGAACGTAATCGAAGAGGGGGTA 120
Qy      121  ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAA 180
Db      121  ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAA 180
Qy      181  GACCTTGGCGCTTTTGGACCGCGCGATGTCTGATTAGTAGTTGGTGGGTTAAAGGCTTAC 240
Db      181  GACCTTGGCGCTTTTGGACCGCGCGATGTCTGATTAGTAGTTGGTGGGTTAAAGGCTTAC 240
Qy      241  CAAGCGACGATCAATATGCTGTGAGAGAGCAGACGCACTAGGGAATGAGACACG 300
Db      241  CAAGCGACGATCAATATGCTGTGAGAGAGCAGACGCACTAGGGAATGAGACACG 300
Qy      301  GCCCAGACTCTTACCGGAGGCGACAGTGGGGAATTTTGGACAAATGGGCGCAACCTGTATC 360
Db      301  GCCCAGACTCTTACCGGAGGCGACAGTGGGGAATTTTGGACAAATGGGCGCAACCTGTATC 360
Qy      361  CAGCAATCCCGGTGATGGAAGAGCCTTGGGTTGTTAAAGCTCTTTCAAGTCAGGAAGA 420
Db      361  CAGCAATCCCGGTGATGGAAGAGCCTTGGGTTGTTAAAGCTCTTTCAAGTCAGGAAGA 420
Qy      421  AAAAGTTACGGTAATTAATCGTGAACCCATGACGGTATGACAGAAAGACACCGGCTAAC 480
Db      421  AAAAGTTACGGTAATTAATCGTGAACCCATGACGGTATGACAGAAAGACACCGGCTAAC 480
Qy      481  TACGTGCAGACGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGGAATTAAGTGGGCGT 540
Db      481  TACGTGCAGACGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGGAATTAAGTGGGCGT 540
Qy      541  AAAAGGTGCGCAGCGGCGCTTGTAAAGTGAATTCCTCGGGCTTTAACTGGGAAT 600
Db      541  AAAAGGTGCGCAGCGGCGCTTGTAAAGTGAATTCCTCGGGCTTTAACTGGGAAT 600
Qy      601  TGGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660
Db      601  TGGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660
Qy      661  TGAATGCGTAGAGATATGAAAGACATCGATGCGCAAGCGACCTCTGGGTTAACT 720
Db      661  TGAATGCGTAGAGATATGAAAGACATCGATGCGCAAGCGACCTCTGGGTTAACT 720
Qy      721  GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATTCCTGGTATGTCACGCC 780
Db      721  GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATTCCTGGTATGTCACGCC 780
```

QY 781 CTTAAAGATGTCATAGTGTGTGGGCTTATAGGCTTGTAACGAGCTAACGGGTGA 840  
Db 781 CTTAAAGATGTCATAGTGTGTGGGCTTATAGGCTTGTAACGAGCTAACGGGTGA 840  
QY 841 AGTTGACCGGCTGGGGAGTACGGTCCGAAAGATTAAATCAAGAAATTGACGGGGACCC 900  
Db 841 AGTTGACCGGCTGGGGAGTACGGTCCGAAAGATTAAATCAAGAAATTGACGGGGACCC 900  
QY 901 GCAACAGCGGTGATATATGATGATTATTCGATGCAACGGGAAACCTTACCTACCTT 960  
Db 901 GCAACAGCGGTGATATATGATGATTATTCGATGCAACGGGAAACCTTACCTACCTT 960  
QY 961 GACATGATGCAATTTTCTAGAGATAGATTAGTCTTCGGGAAACGGTAAACAGAGTCTG 1020  
Db 961 GACATGATGCAATTTTCTAGAGATAGATTAGTCTTCGGGAAACGGTAAACAGAGTCTG 1020  
QY 1021 CATGGCTGTGTCAGCTCGTGTCTGAGATGTTGGGTTAAAGTCCCGAACGAGCGCAAC 1080  
Db 1021 CATGGCTGTGTCAGCTCGTGTCTGAGATGTTGGGTTAAAGTCCCGAACGAGCGCAAC 1080  
QY 1081 CTTGTCAATTAATGTCATCAATTTGTTGGGCACTTATGAACTGCGGGTGAACACCG 1140  
Db 1081 CTTGTCAATTAATGTCATCAATTTGTTGGGCACTTATGAACTGCGGGTGAACACCG 1140  
QY 1141 GAGGAAGTGGGAGTGAAGTCAAGTCTCATGGCCCTTATGAGGTAGGGCTTCAACGTAA 1200  
Db 1141 GAGGAAGTGGGAGTGAAGTCAAGTCTCATGGCCCTTATGAGGTAGGGCTTCAACGTAA 1200  
QY 1201 TACAAATGGCGCGTACAGAGGGTTGCGAACCCGCGAGGGGAGCTAATCTCAAGAAAGCGG 1260  
Db 1201 TACAAATGGCGCGTACAGAGGGTTGCGAACCCGCGAGGGGAGCTAATCTCAAGAAAGCGG 1260  
QY 1261 TCGTAATGTCGAGTCCGAGTCTGCAACTGCTCGTGAAGTCGGAATCGTAAGTATCGC 1320  
Db 1261 TCGTAATGTCGAGTCCGAGTCTGCAACTGCTCGTGAAGTCGGAATCGTAAGTATCGC 1320  
QY 1321 GGATCAGCATGTGCGGGTGAATAGTTCGCGGGTCTTGTACACACCGCCGTCACACAT 1380  
Db 1321 GGATCAGCATGTGCGGGTGAATAGTTCGCGGGTCTTGTACACACCGCCGTCACACAT 1380  
QY 1381 GGGAGTGGGTTTCAACGAAAGCAGTATGTTAAACGTAAGAGGGCCCTTGCCACGGTGA 1440  
Db 1381 GGGAGTGGGTTTCAACGAAAGCAGTATGTTAAACGTAAGAGGGCCCTTGCCACGGTGA 1440  
QY 1441 GATTCAATGACTGGGGT 1457  
Db 1441 GATTCAATGACTGGGGT 1457

RESULT 7  
LOCUS AX316092 1457 bp DNA linear PAT 14-DEC-2001  
DEFINITION Sequence 1 from Patent WO0190312.  
ACCESSION AX316092  
VERSION AX316092.1 GI:17899283  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
1 unclassified sequences.

REFERENCE  
AUTHORS Hovanec T.A. and Burrell, P.C.  
TITLE Ammonia-oxidizing bacteria  
JOURNAL Patent: WO 0190312-A 1 29-NOV-2001;  
ADUARIA, INC. (US)  
FEATURES  
source Location/Qualifiers  
1. 1457  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="A08 Type A ammonia-oxidizing bacterium represented  
by R7 clone140"

ORIGIN

Query Match 99.7%; Score 1452.2; DB 6; Length 1457;  
Best Local Similarity 99.8%; Pred. No. 9.7e-16;  
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATTTGAACGCTGGCGGATCTTTTACATGATGAAGTCGAACGGCAGACCGATGCTTGCAT 60  
Db 1 ATTTGAACGCTGGCGGATCTTTTACATGATGAAGTCGAACGGCAGACCGATGCTTGCAT 60  
QY 61 CTGTGGCAGTGGCGGAGCGGGTGAATGATCGGAACGTATCCAAAGAGGGGGTGA 120  
Db 61 CTGTGGCAGTGGCGGAGCGGGTGAATGATCGGAACGTATCCAAAGAGGGGGTGA 120  
QY 121 AGCCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAAAGCAGGGGATCGAA 180  
Db 121 AGCCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAAAGCAGGGGATCGAA 180  
QY 181 GACCTTGGGCTTTTGAAGGGCGGATGTCTGATTAGCTAGTTGTTGGGGTAAAGGCTTAC 240  
Db 181 GACCTTGGGCTTTTGAAGGGCGGATGTCTGATTAGCTAGTTGTTGGGGTAAAGGCTTAC 240  
QY 241 CAAGCGAGATCAGTATGTTGGTCTGAGAGAGACGACCACTTGGATCTGAGACAGC 300  
Db 241 CAAGCGAGATCAGTATGTTGGTCTGAGAGAGACGACCACTTGGATCTGAGACAGC 300  
QY 301 GCCCAACTCTTACGGGAGGACGACAGTGGGGAATTTTGAACAATGGGCCCAAGCTGATC 360  
Db 301 GCCCAACTCTTACGGGAGGACGACAGTGGGGAATTTTGAACAATGGGCCCAAGCTGATC 360  
QY 361 CAGCAATGCGCGGTGATGAAGAGGCTTGGGGTTGAAGGCTCTTTCAGTGAAGA 420  
Db 361 CAGCAATGCGCGGTGATGAAGAGGCTTGGGGTTGAAGGCTCTTTCAGTGAAGA 420  
QY 421 AAAGTTAAGGTAAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 480  
Db 421 AAAGTTAAGGTAAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 480  
QY 481 TACGTCCACAGACCGCGGTGAATAGTGAAGTGAATGTAATGTAATGTAATGTAATGTAATG 540  
Db 481 TACGTCCACAGACCGCGGTGAATAGTGAAGTGAATGTAATGTAATGTAATGTAATGTAATG 540  
QY 541 AAAGGTGGCGCAGGCGGCTTGTAGTCAATGTAATGTAATGTAATGTAATGTAATGTAATG 600  
Db 541 AAAGGTGGCGCAGGCGGCTTGTAGTCAATGTAATGTAATGTAATGTAATGTAATGTAATG 600  
QY 601 AAAGGTGGCGCAGGCGGCTTGTAGTCAATGTAATGTAATGTAATGTAATGTAATGTAATG 660  
Db 601 AAAGGTGGCGCAGGCGGCTTGTAGTCAATGTAATGTAATGTAATGTAATGTAATGTAATG 660  
QY 661 TGAATGCGTGAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 720  
Db 661 TGAATGCGTGAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 720  
QY 721 GAGGCTATGCAAGAAAGCTGGGGAGCAACAGAAATTGATACCTGCTGTAAGTCAAGCC 780  
Db 721 GAGGCTATGCAAGAAAGCTGGGGAGCAACAGAAATTGATACCTGCTGTAAGTCAAGCC 780  
QY 781 CTTAAAGATGTCATAGTGTGTGGGCTTATAGGCTTGTAACGAGCTAACGGGTGA 840  
Db 781 CTTAAAGATGTCATAGTGTGTGGGCTTATAGGCTTGTAACGAGCTAACGGGTGA 840  
QY 841 AGTTGACCGGCTGGGGAGTACGGTCCGAAAGATTAAATCAAGAAATTGACGGGGACCC 900  
Db 841 AGTTGACCGGCTGGGGAGTACGGTCCGAAAGATTAAATCAAGAAATTGACGGGGACCC 900  
QY 901 GCAACAGCGGTGATATATGATGATTATTCGATGCAACGGGAAACCTTACCTACCTT 960  
Db 901 GCAACAGCGGTGATATATGATGATTATTCGATGCAACGGGAAACCTTACCTACCTT 960  
QY 961 GACATGATGCAATTTTCTAGAGATAGATTAGTCTTCGGGAAACGGTAAACAGAGTCTG 1020  
Db 961 GACATGATGCAATTTTCTAGAGATAGATTAGTCTTCGGGAAACGGTAAACAGAGTCTG 1020

Qy	1021	CATGGCTGTCAGAGCTCGGTGCTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACC	1080
Db	1021	CATGGCTGTCAGAGCTCGGTGCTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACC	1080
Qy	1081	CTTGTCATTAAATTTGCCATATTTGGTGGGCACTTTAATGAGACTGCGGGTGA CAACCG	1140
Db	1081	CTTGTCATTAAATTTGCCATATTTGGTGGGCACTTTAATGAGACTGCGGGTGA CAACCG	1140
Qy	1141	GAGGAAGGTGGGGATATACGTCAAGTCTCTCATGGCCCTTTATGGGTAGGGCTTCAACGTAA	1200
Db	1141	GAGGAAGGTGGGGATATACGTCAAGTCTCTCATGGCCCTTTATGGGTAGGGCTTCAACGTAA	1200
Qy	1201	TACCAATGCGCGGATACAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAGAAAGCGCG	1260
Db	1201	TACCAATGCGCGGATACAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAGAAAGCGCG	1260
Qy	1261	TGCTAGTCCGGATCGGAGTCTGCAACTCGACTCCGTTGAAGTCGGAATGCTAGTAATCGC	1320
Db	1261	TGCTAGTCCGGATCGGAGTCTGCAACTCGACTCCGTTGAAGTCGGAATGCTAGTAATCGC	1320
Qy	1321	GGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTTGTAACAACCGCCCGTCAACCAT	1380
Db	1321	GGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTTGTAACAACCGCCCGTCAACCAT	1380
Qy	1381	GGAGTGGGTTTACACGAGAGCAGTAGTCTAAACCGTAAGAGAGGCGCTTGCCACGCTGA	1440
Db	1381	GGAGTGGGTTTACACGAGAGCAGTAGTCTAAACCGTAAGAGAGGCGCTTGCCACGCTGA	1440
Qy	1441	GATTCATGACTGGGGTG 1457	
Db	1441	GATTCATGACTGGGGTG 1457	

RESULT 8  
AX316093 1457 bp DNA linear PAT 14-DEC-2001  
LOCUS AX316093 Sequence 2 from Patent WO0190312.  
DEFINITION AX316093  
ACCESSION AX316093  
VERSION AX316093.1 GI:17899284  
KEYWORDS  
ORGANISM  
SOURCE  
REFERENCE  
1  
unidentified  
unclassified  
unclassified sequences.

AUTHORS Hovanec, T. A. and Burrell, P. C.  
TITLE Ammonia-oxidizing bacteria  
JOURNAL Patent: WO 0190312-A 2 29-NOV-2001;  
AQUARIA, INC. (US)  
FEATURES  
source 1. 1457  
location/Qualifiers  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="AOB Type A1 ammonia-oxidizing bacterium represented  
by R7 clone187"

ORIGIN  
Query Match 99.7%; Score 1452.2; DB 6; Length 1457;  
Best Local Similarity 99.8%; Pred. No. 9.7e-16;  
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	ATTGAACGCTGGCGGATGCTTTACATGCAAGTGAACGGGACAGAGGATGCTTGCAAT	60
Db	1	ATTGAACGCTGGCGGATGCTTTACATGCAAGTGAACGGGACAGAGGATGCTTGCAAT	60
Qy	61	CTGTGTGCGAGATGCGGAGACGGGTGAGTAATGATCGGAACGTATCCAGAAAGAGGGGGTA	120
Db	61	CTGTGTGCGAGATGCGGAGACGGGTGAGTAATGATCGGAACGTATCCAGAAAGAGGGGGTA	120
Qy	121	ACGCATCGAAAGATGTCTAATACCGCATATATCTTAAAGAGAGAAAGCAGGGGATCGAAA	180
Db	121	ACGCATCGAAAGATGTCTAATACCGCATATATCTTAAAGAGAGAAAGCAGGGGATCGAAA	180

Qy	181	GACCTTGCGCTTTTGAAGGCGCGGATGTCATAGTACAGTTGTTGGGTAAAGGCTTAC	240
Db	181	GACCTTGCGCTTTTGAAGGCGCGGATGTCATAGTACAGTTGTTGGGTAAAGGCTTAC	240
Qy	241	CAAGCGACGATCACTAGTGTGCTGAGAGAGCAACCACTCTGGACTGAGACACG	300
Db	241	CAAGCGACGATCACTAGTGTGCTGAGAGAGCAACCACTCTGGACTGAGACACG	300
Qy	301	GCCCGACTCCCTACCGGAGGCGAGCGTGGGAAATTTGACAAATGGGGCGCAAGCTGATC	360
Db	301	GCCCGACTCCCTACCGGAGGCGAGCGTGGGAAATTTGACAAATGGGGCGCAAGCTGATC	360
Qy	361	CAGCAATCCCGGTGAGTGAAGAAAGCCCTCGGGTTGTAAGCTCTTTCAGTCGAGAA	420
Db	361	CAGCAATCCCGGTGAGTGAAGAAAGCCCTCGGGTTGTAAGCTCTTTCAGTCGAGAA	420
Qy	421	AAAGGTTACGGTAAATTCGTGACCATGACGATTCGACAGAAAGAACACCGGCTAAC	480
Db	421	AAAGGTTACGGTAAATTCGTGACCATGACGATTCGACAGAAAGAACACCGGCTAAC	480
Qy	481	TACGTGCGACGACCGCGGTTAATACGTAGGGGTCAGAGGGTTAATGGAATTAATCTGGGGCT	540
Db	481	TACGTGCGACGACCGCGGTTAATACGTAGGGGTCAGAGGGTTAATGGAATTAATCTGGGGCT	540
Qy	541	AAAGGTCGCGACGCGGCTTGTAAAGTGAATGTAATCCCGGGCTTTAACTTGGGAAT	600
Db	541	AAAGGTCGCGACGCGGCTTGTAAAGTGAATGTAATCCCGGGCTTTAACTTGGGAAT	600
Qy	601	TGCGTTTGAACCTACAAAGCTAGTAGTGGGACAGAGGAGTGGAAATTCATGTGTAGCAG	660
Db	601	TGCGTTTGAACCTACAAAGCTAGTAGTGGGACAGAGGAGTGGAAATTCATGTGTAGCAG	660
Qy	661	TGAATGGGTGAGATATGGAAGAACATCGATGGGAGGACGCTCCGGGTTAACT	720
Db	661	TGAATGGGTGAGATATGGAAGAACATCGATGGGAGGACGCTCCGGGTTAACT	720
Qy	721	GACGCTCATGACGAAACGTCGGGAGCAACAGAGATTAGTACCTGTAGTCCACGCC	780
Db	721	GACGCTCATGACGAAACGTCGGGAGCAACAGAGATTAGTACCTGTAGTCCACGCC	780
Qy	781	CTAAACGATGTCACCTAGTGTGGGCTTTATAGGCTTGGTAAACGAAAGCTAACCCGTGA	840
Db	781	CTAAACGATGTCACCTAGTGTGGGCTTTATAGGCTTGGTAAACGAAAGCTAACCCGTGA	840
Qy	841	AGTTGACCGGCTGGGAGATGCGTGCAGAAATTAACTCAAGAAATTTGACGGGACCC	900
Db	841	AGTTGACCGGCTGGGAGATGCGTGCAGAAATTAACTCAAGAAATTTGACGGGACCC	900
Qy	901	GCACAAAGCGTGGATTAATGATGATTAATTCATGCAACGCGAAAAACCTTACTTACCTT	960
Db	901	GCACAAAGCGTGGATTAATGATGATTAATTCATGCAACGCGAAAAACCTTACTTACCTT	960
Qy	961	GACATGTACGAATTTTCTAGAGATGATGATGCTTGGGAAACGCTTAACACAGGTGCTG	1020
Db	961	GACATGTACGAATTTTCTAGAGATGATGATGCTTGGGAAACGCTTAACACAGGTGCTG	1020
Qy	1021	CATGCTGTGCTGACGCTGTCGTGCGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACC	1080
Db	1021	CATGCTGTGCTGACGCTGTCGTGCGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACC	1080
Qy	1081	CTTGTCATTAAATTTGCCATATTTGGTGGGCACTTTAATGAGACTGCGGGTGA CAACCG	1140
Db	1081	CTTGTCATTAAATTTGCCATATTTGGTGGGCACTTTAATGAGACTGCGGGTGA CAACCG	1140
Qy	1141	GAGGAAGGTGGGGATGAGTCAAGTCTCATAGGCCCTTTATGGGTAGGGCTTCAACGTAA	1200
Db	1141	GAGGAAGGTGGGGATGAGTCAAGTCTCATAGGCCCTTTATGGGTAGGGCTTCAACGTAA	1200
Qy	1201	TACCAATGCGCGTACAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAGAAAGCGCG	1260
Db	1201	TACCAATGCGCGTACAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAGAAAGCGCG	1260
Qy	1261	TGCTAGTCCGGATCGGAATCTGCAACTCGACTCCGTTGAAGTCGGAATGCTAGTAATCGC	1320

|||||  
Db 1261 TCGTAGTCCGAGTCGAGTCTGCACTCCGTAAGTCGGATCCCTGATATCCG 1320  
|||  
Qy 1321 GGATCGCATGTCGGGGTGAATACGTTCCGGGCTTTGTACACACCGCCGCTCACCAT 1380  
|||  
Db 1321 GGATCGCATGTCGGGGTGAATACGTTCCGGGCTTTGTACACACCGCCGCTCACCAT 1380  
|||  
Qy 1381 GGGAGTGGGTTTCCACGAGAGCAGTATCTAACCGTAAAGAGGGCGCTTGCCACGGTA 1440  
|||  
Db 1381 GGGAGTGGGTTTCCACGAGAGCAGTATCTAACCGTAAAGAGGGCGCTTGCCACGGTA 1440  
|||  
Qy 1441 GATTCAATGACTGGGGTG 1457  
|||  
Db 1441 GATTCAATGACTGGGGTG 1457  
|||  
RESULT 9  
AF386752 1457 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosomonas sp. R7c131 16S ribosomal RNA gene, partial sequence.  
DEFINITION AP386752  
ACCESSION AP386752  
VERSION AP386752.1 GI:117864829  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. R7c131  
ORGANISM Nitrosomonas sp. R7c131  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas; environmental samples.  
REFERENCE 1 (bases 1 to 1457)  
Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
1172936  
2 (bases 1 to 1457)  
Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
Direct Submision  
Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condon Dr., Moorpark, CA 93021, USA  
LOCATION/Qualifiers  
FEATURES  
source  
1..1457  
/organism="Nitrosomonas sp. R7c131"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichments"  
/db\_xref="taxon:180511"  
/clone="R7c131"  
/environmental\_sample  
<1..>1457  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 99.6%; Score 1450.6; DB 3; Length 1457;  
Best Local Similarity 99.7%; Pred. No. 1e-15;  
Matches 1453; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATTTGAAGCTGGCGGACATGCTTTACACATGCAATCGAACCGGACGACGATGCTTGAT 60  
|||  
Db 1 ATTTGAAGCTGGCGGACATGCTTTACACATGCAATCGAACCGGACGACGATGCTTGAT 60  
|||  
Qy 61 CTGGTGGCGAGTGGCGGACGGGTGATATGATCGAACGATTCAGAAAGAGGGGGTA 120  
|||  
Db 61 CTGGTGGCGAGTGGCGGACGGGTGATATGATCGAACGATTCAGAAAGAGGGGGTA 120  
|||  
Qy 121 ACGCATGAAAGATGTGCTAATACCGCATATACCTTAAGAGAGAAAGAGGGGATCGAAA 180  
|||  
Db 121 ACGCATGAAAGATGTGCTAATACCGCATATACCTTAAGAGAGAAAGAGGGGATCGAAA 180  
|||  
Qy 181 GACCTTGGGCTTTTGGAGCGGCGGATGTCTGATTAGCTAGTTGGTGGGGTAAAGCCCTAC 240  
|||  
Db 181 GACCTTGGGCTTTTGGAGCGGCGGATGTCTGATTAGCTAGTTGGTGGGGTAAAGCCCTAC 240  
|||  
Qy 241 CAAGGCGACGATCAGTACTGTGCTGAGAGACGACGACCACTGGGACTGAGACACG 300  
|||

Db 241 CAAGGCGACGATCAGTACTGTGCTGAGAGACGACGACCACTGGGACTGAGACACG 300  
|||  
Qy 301 GCCCAAGCTCCTACCGGAGCGCAGAGTGGGAAATTTTGGACAAATGGGGCGAAAGCTGATC 360  
|||  
Db 301 GCCCAAGCTCCTACCGGAGCGCAGAGTGGGAAATTTTGGACAAATGGGGCGAAAGCTGATC 360  
|||  
Qy 361 CAGCAATGCGGGTGAAGTGAAGAAAGCCCTCGGGTTGTTAAAGCTCTTTCACTCGAAGAA 420  
|||  
Db 361 CAGCAATGCGGGTGAAGTGAAGAAAGCCCTCGGGTTGTTAAAGCTCTTTCACTCGAAGAA 420  
|||  
Qy 421 AAAGGTTACGATTAATTAATCGTGAACCAATGACGATTCGACGAAAGAACACCGGCTAAC 480  
|||  
Db 421 AAAGGTTACGATTAATTAATCGTGAACCAATGACGATTCGACGAAAGAACACCGGCTAAC 480  
|||  
Qy 481 TACGTGCCAGCAGCCCGGTAATACGTAGGGGTGCAAGCTTAAATCGGAATTAATCTGGGCGT 540  
|||  
Db 481 TACGTGCCAGCAGCCCGGTAATACGTAGGGGTGCAAGCTTAAATCGGAATTAATCTGGGCGT 540  
|||  
Qy 541 AAAGGTTGCGCAGGGGGGCTTGAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAAT 600  
|||  
Db 541 AAAGGTTGCGCAGGGGGGCTTGAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAAT 600  
|||  
Qy 601 TCGCTTTGAAACTACAAAGCTAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660  
|||  
Db 601 TCGCTTTGAAACTACAAAGCTAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660  
|||  
Qy 661 TGAATGCGTGAAGATATGAAAGAAATGATGTGGGAGGACGAGCTCTCGGGTTAACCT 720  
|||  
Db 661 TGAATGCGTGAAGATATGAAAGAAATGATGTGGGAGGACGAGCTCTCGGGTTAACCT 720  
|||  
Qy 721 GACGCTCAGCAGCAAGAAAGGTGGGGAGCAACAGATTGATACCTGTAGTCCAGCC 780  
|||  
Db 721 GACGCTCAGCAGCAAGAAAGGTGGGGAGCAACAGATTGATACCTGTAGTCCAGCC 780  
|||  
Qy 781 CTAAACGATGTCAACTAGTTGTGGGCTTAATTAAGCTTGTATACGAAGCTAACCGGTGA 840  
|||  
Db 781 CTAAACGATGTCAACTAGTTGTGGGCTTAATTAAGCTTGTATACGAAGCTAACCGGTGA 840  
|||  
Qy 841 AGTTGACCGCTGGGGAGTACGGTGCAGAGATTAAATCTCAAGAAATGACGGGAGCC 900  
|||  
Db 841 AGTTGACCGCTGGGGAGTACGGTGCAGAGATTAAATCTCAAGAAATGACGGGAGCC 900  
|||  
Qy 901 GCACAAAGCGGTGATATATGATGATTAATTCGATGCAACGCGAAGAACTTTACTACCTT 960  
|||  
Db 901 GCACAAAGCGGTGATATATGATGATTAATTCGATGCAACGCGAAGAACTTTACTACCTT 960  
|||  
Qy 961 GACATGAGCGAATTTTCTAGAGATAGATTAGTCTTCGGGAAACGTTAACACAGGTGCTG 1020  
|||  
Db 961 GACATGAGCGAATTTTCTAGAGATAGATTAGTCTTCGGGAAACGTTAACACAGGTGCTG 1020  
|||  
Qy 1021 CATGGCTGTGTCAGCTCGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCCGCAAC 1080  
|||  
Db 1021 CATGGCTGTGTCAGCTCGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCCGCAAC 1080  
|||  
Qy 1081 CTTGTCAATTAATGCCATCATTTTGGTGGGCACTTTAATGAGACTGCCGTGCAAAACG 1140  
|||  
Db 1081 CTTGTCAATTAATGCCATCATTTTGGTGGGCACTTTAATGAGACTGCCGTGCAAAACG 1140  
|||  
Qy 1141 GAGGAAGGTGGGAGTGAAGTCAAGTCCATGAGCCCTTAATGGGTAGGGCTTCAACGTTAA 1200  
|||  
Db 1141 GAGGAAGGTGGGAGTGAAGTCAAGTCCATGAGCCCTTAATGGGTAGGGCTTCAACGTTAA 1200  
|||  
Qy 1201 TACAAATGGCGGTACAGAGGGTTGCCAACCCGAGGGGGAGCTAATCTCAAGAAAGCGC 1260  
|||  
Db 1201 TACAAATGGCGGTACAGAGGGTTGCCAACCCGAGGGGGAGCTAATCTCAAGAAAGCGC 1260  
|||  
Qy 1261 TCGTAGTCGGAGTCCGAGTTCGCAATCGCATCCGGAAGATCGTAAGTAATCGC 1320  
|||  
Db 1261 TCGTAGTCGGAGTCCGAGTTCGCAATCGCATCCGGAAGATCGTAAGTAATCGC 1320  
|||  
Qy 1321 GGATCAGCATGTGCGGGTGAATACGTTCCGGGGTCTTTGTACACACCGCCGCTCACACAT 1380  
|||  
Db 1321 GGATCAGCATGTGCGGGTGAATACGTTCCGGGGTCTTTGTACACACCGCCGCTCACACAT 1380  
|||



RESULT 10  
AF386751 1426 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosomonas sp. R7c155 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386751  
ACCESSION AF386751 GI:17864828  
VERSION AF386751.1 GI:17864828  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. R7c155  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosomonas; environmental samples.  
REFERENCE 1 (bases 1 to 1426)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 11722936  
REFERENCES 2 (bases 1 to 1426)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria Group, 6100 Condon Dr. Moorpark, CA 93021, USA  
Location/Qualifiers  
1..1426  
/organism="Nitrosomonas sp. R7c155"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated from freshwater aquaria enrichments"  
/db\_xref="taxon:180510"  
/clone="R7c155"  
/environmental\_sample  
<1..>1426  
/product="16S ribosomal RNA"  
rRNA  
ORIGIN  
Query Match 97.6%; Score 1421.8; DB 3; Length 1426;  
Best Local Similarity 99.8%; Pred. No. 3e-15;  
Matches 1423; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 32 AGTGAACGGGACGACCGGATCTTGATCTGTGCGGACGCGGACGGGTGAAGTAA 91  
DB 1 AAGTGAAAGGACGACCGGATCTTGATCTGTGCGGACGCGGACGGGTGAAGTAA 60  
QY 92 CATCGAAGCTATCCAGAAAGAGGGGGTAAAGCATGAAAGATGTCTAATACCGCATAT 151  
DB 61 CATCGAAGCTATCCAGAAAGAGGGGGTAAAGCATGAAAGATGTCTAATACCGCATAT 120  
QY 152 ACTTAAAGAGGAAAGACAGGGGATGAAAGACCTTGCGCTTTTGAAGCGGCGGATGTCTG 211  
DB 121 ACTTAAAGAGGAAAGACAGGGGATGAAAGACCTTGCGCTTTTGAAGCGGCGGATGTCTG 180  
QY 212 ATTACTAGTGTGTGGGTAAAGGCTTACCAAGCGGACGATCAGTATGTGTGAGAG 271  
DB 181 ATTACTAGTGTGTGGGTAAAGGCTTACCAAGCGGACGATCAGTATGTGTGAGAG 240  
QY 272 ACGACGAGCACACTGGGACTGAGACACGGCCAGACTCTTAACGGAGGACAGAGTGGG 331  
DB 241 ACGACGAGCACACTGGGACTGAGACACGGCCAGACTCTTAACGGAGGACAGAGTGGG 300  
QY 332 AATTTTGAACAATGGGCGGACCTGATCAGCAATGCCGCTGAGTGAAGAGGCTTTC 391  
DB 301 AATTTTGAACAATGGGCGGACCTGATCAGCAATGCCGCTGAGTGAAGAGGCTTTC 360

QY 392 GGGTTGAAGCTCTTTCAGTCGAGAGAAAAGTTACGGTAAATATTCGATCCATGA 451  
DB 361 GGGTTGAAGCTCTTTCAGTCGAGAGAAAAGTTACGGTAAATATTCGATCCATGA 420  
QY 452 CGGTATCGACAGAAAGACACCGGCTAACTACGTGCCAGACCGCGGTAAATACGAGG 511  
DB 421 CGGTATCGACAGAAAGACACCGGCTAACTACGTGCCAGACCGCGGTAAATACGAGG 480  
QY 512 TGCAGCGTTATTCGGAATTAATGCGCGTAAAGGGTGGCGGCGGCTTGAATCGA 571  
DB 481 TGCAGCGTTATTCGGAATTAATGCGCGTAAAGGGTGGCGGCGGCTTGAATCGA 540  
QY 572 TGTGAATTCGCCGGGCTTAACCTGGGAAATGCGTTGAAAACCTACAAAGCTAGAGTGGC 631  
DB 541 TGTGAATTCGCCGGGCTTAACCTGGGAAATGCGTTGAAAACCTACAAAGCTAGAGTGGC 600  
QY 632 AGAGGAGGTGGAATTCATGTGTAGCAAGTGAATGCGTAGAATATGAGAAACATCGA 691  
DB 601 AGAGGAGGTGGAATTCATGTGTAGCAAGTGAATGCGTAGAATATGAGAAACATCGA 660  
QY 692 TGGCGAAGGACCTCTCGGCTTAACATGACGCTCATGCAAGAAAGCTGGGGAGCAAA 751  
DB 661 TGGCGAAGGACCTCTCGGCTTAACATGACGCTCATGCAAGAAAGCTGGGGAGCAAA 720  
QY 752 CAGATTTAGATACCTGGTATGTCACGCGCTTAAAGATGTCATAGTTGTTGGCCTTA 811  
DB 721 CAGATTTAGATACCTGGTATGTCACGCGCTTAAAGATGTCATAGTTGTTGGCCTTA 780  
QY 812 TTAGGCTTGTATACGAGCTAAACGCGTGAAGTTGACCGCTTGGGAGTACGGTCCAAAGA 871  
DB 781 TTAGGCTTGTATACGAGCTAAACGCGTGAAGTTGACCGCTTGGGAGTACGGTCCAAAGA 840  
QY 872 TTAATACTCAAGGAATTTGACGGGACCCGCAAGCGGTGATATGATTAATTCG 931  
DB 841 TTAATACTCAAGGAATTTGACGGGACCCGCAAGCGGTGATATGATTAATTCG 900  
QY 932 ATGCAACCGGAAAACCTTACCTTACCTTGAATGATGCAATTTTCTAGAGATGATTA 991  
DB 901 ATGCAACCGGAAAACCTTACCTTACCTTGAATGATGCAATTTTCTAGAGATGATTA 960  
QY 992 GTGCTTGGGAAACGCTTAAACAGGTGTCATGCGTGTGTCAGCTGCTGTCGAGATG 1051  
DB 961 GTGCTTGGGAAACGCTTAAACAGGTGTCATGCGTGTGTCAGCTGCTGTCGAGATG 1020  
QY 1052 TTGGGTAAAGTCCCGCAACGAGCGCAACCTTGTCAATTAATGCCATATTTGTTGGC 1111  
DB 1021 TTGGGTAAAGTCCCGCAACGAGCGCAACCTTGTCAATTAATTTGTTGGG 1080  
QY 1112 ACTTAAATGAGACTCCGGTGAACAACGAGAAAGTGGGGATGACGTCAGAGTCTCAT 1171  
DB 1081 ACTTAAATGAGACTCCGGTGAACAACGAGAAAGTGGGGATGACGTCAGAGTCTCAT 1140  
QY 1172 GGCCTTATGGGTAGGGCTTCAACAGTAAATCAATGGCGCTTACAGAGGGTTGCCAACCC 1231  
DB 1141 GGCCTTATGGGTAGGGCTTCAACAGTAAATCAATGGCGCTTACAGAGGGTTGCCAACCC 1200  
QY 1232 GCGAGGGGAGTAAATCTCAAGAAAGCGGTGTATCCGGATCGGAATCGGAATCTGAC 1291  
DB 1201 GCGAGGGGAGTAAATCTCAAGAAAGCGGTGTATCCGGATCGGAATCTGAC 1260  
QY 1292 TCCGTGAAGTGGGAATTCGCTAGTAATTCGCGGATCGAGATGTCGCGGTAAATCGTTCCG 1351  
DB 1261 TCCGTGAAGTGGGAATTCGCTAGTAATTCGCGGATCGAGATGTCGCGGTAAATCGTTCCG 1320  
QY 1352 GGTCTTGTACACACCGCCGTCACACATGGAGTGGGTTTCCAGAGACGATAGTCT 1411  
DB 1321 GGTCTTGTACACACCGCCGTCACACATGGAGTGGGTTTCCAGAGACGATAGTCT 1380  
QY 1412 AACGTTAAGAGGGGCTTGGCCAGGTGAGATTCATGACTGGGGTG 1457  
DB 1381 AACGTTAAGAGGGGCTTGGCCAGGTGAGATTCATGACTGGGGTG 1426



RESULT 11  
AF386746 1487 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosomonas sp. BF16c57 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386746  
ACCESSION AF386746.1 GI:17864823  
VERSION  
KEYWORDS ENV.  
SOURCE Nitrosomonas sp. BF16c57  
ORGANISM Nitrosomonas sp. BF16c57  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas; environmental samples.  
REFERENCE  
AUTHORS 1 (bases 1 to 1487)  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 1172936  
REFERENCE 2 (bases 1 to 1487)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condon Dr. Moorpark, CA 93021, USA  
FEATURES  
source  
1. .1487  
Location/Qualifiers  
/organism="Nitrosomonas sp. BF16c57"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichments"  
/db\_xref="taxon:180505"  
/clone="BF16c57"  
/environmental\_sample  
<1. .1487  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 95.4%; Score 1390; DB 3; Length 1487;  
Best Local Similarity 97.5%; Pred. No. 8.9e-15;  
Matches 1422; Conservative 0; Mismatches 35; Indels 1; Gaps 1;  
QY 1 ATTGAAGCTGGCGGCGGCTTTACATGCAAGTCGAAACGGACGACGCGATGCTTGCAT 60  
DB 18 ATTGAAGCTGGCGGCGGCTTTACATGCAAGTCGAAACGGACGACGCGGATGCTTGCAC 77  
QY 61 CTGGTGGCGATGGCGGACGGGTAGTATGATCGAACGCTATCCAGAAAGGGGGTAT 120  
DB 78 CTGGTGGCGATGGCGGACGGGTAGTATGATCGAACGCTATCCAGAAAGGGGGTAT 137  
QY 121 ACGCATGGAAGATGTCTAATACCGCATATCTCTAAGAGAAAGCAGGGGATCGAAA 180  
DB 138 ACGCATGGAAGATGTCTAATACCGCATATCTCTAAGAGAAAGCAGGGGATCGAAA 197  
QY 181 GACCTTGGCGTTTGGAGCGGCCGATGCTGTATTAAGTATGGTGGGTAAAGGCTTAC 240  
DB 198 GACCTTGGCGTTTGGAGCGGCCGATGCTGTATTAAGTATGGTGGGTAAAGGCTTAC 257  
QY 241 CAAAGCGCATCGATAGTGTCTGAGAGAGACGACCACTGGGACTGAGACACG 300  
DB 258 CAAAGCGCATCGATAGTGTCTGAGAGAGACGACCACTGGGACTGAGACACG 317  
QY 301 GCCCAAGCTCTTAACGGAGGACGAGTGGGGAATTTTGAACATATGGGCGCAAGCCTGATC 360  
DB 318 GCCCAAGCTCTTAACGGAGGACGAGTGGGGAATTTTGAACATATGGGCGCAAGCCTGATC 377  
QY 361 CAGCAATGCCCGGTGAGTGAAGAAAGGCTTGGGTTGTAAGCTTTTCAATCGAGAA 420  
DB 378 CAGCAATGCCCGGTGAGTGAAGAAAGGCTTGGGTTGTAAGCTTTTCAATCGAGAA 437  
QY 421 AAAAGTTACGGTAATATCGTACCATGACGATGACAGAGAAAGACCCGGCTAAC 480  
DB 438 AAAAGTTACGGTAATATCGTACCATGATGACGATGACAGAGAAAGACCCGGCTAAC 497  
QY 481 TACGTGCCAGGACCGCGGTATATGATGGGTGCAAGCGTTTATCGAATTAATCTGGCGCT 540

DB 498 TACGTGCCAGGACCGCGGTATATGATGGGTGCAAGCGTTTATCGAATTAATCTGGCGCT 557  
QY 541 AAAAGTTACGGTAATATCGTACCATGATGACGATGACGAGTGAATCCCGGGCTTAACCTGGGAAT 600  
DB 558 AAAAGTTACGGTAATATCGTACCATGATGACGATGACGAGTGAATCCCGGGCTTAACCTGGGAAT 617  
QY 601 TCGGTTTGAATATCAAAAGCTAGAGTGTGACAGAGGAGTGGAAATTCATGTGTAGCAG 660  
DB 618 TCGGTTTGAATATCAAAAGCTAGAGTGTGACAGAGGAGGAGTGGAAATTCATGTGTAGCAG 677  
QY 661 TGAATATGCTAGAGATATGAAAGAAATCATGATGGCGAAGCAGCTCTGGGTTAACT 720  
DB 678 TGAATATGCTAGAGATATGAAAGAAATCATGATGGCGAAGCAGCTCTGGGTTAACT 737  
QY 721 GACGCTATGACGAAAGCGTGGGAGCAACAGATTTGATATCCCTGGTATGTCACGCG 780  
DB 738 GACGCTATGACGAAAGCGTGGGAGCAACAGATTTGATATCCCTGGTATGTCACGCG 797  
QY 781 CTAAAGATGTCAAATGATTTGTTGGGCTTATTAAGCTTGTAAACGAAAGCTAACGCGTGA 840  
DB 798 CTAAAGATGTCAAATGATTTGTTGGGCTTATTAAGCTTGTAAACGAAAGCTAACGCGTGA 857  
QY 841 AGTTGACCGCTGGGAGTACGGTTCGCAAGATTAAACTCAAGAAATTGACGGGACCC 900  
DB 858 AGTTGACCGCTGGGAGTACGGTTCGCAAGATTAAACTCAAGAAATTGACGGGACCC 917  
QY 901 GCACAAAGCGGTGATATGATGCTTAAATTCGATGCAACCGGAAACCTTAACTTACCTT 960  
DB 918 GCACAAAGCGGTGATATGATGCTTAAATTCGATGCAACCGGAAACCTTAACTTACCTT 977  
QY 961 GACATGTAGCGAATTTTCTAAGATATGATTAAGT-CTTGGGAAACGCTAACACAGGTGCT 1019  
DB 978 GACATGTAGCGAATTTTCTAAGATATGATTAAGT-CTTGGGAAACGCTAACACAGGTGCT 1037  
QY 1020 GCATGCGTGTCTCACTGCTGTCTGTAGATGTTGGTTAAAGTCCCGCAAGCGCGAAC 1079  
DB 1038 GCATGCGTGTCTCACTGCTGTCTGTAGATGTTGGTTAAAGTCCCGCAAGCGCGAAC 1097  
QY 1080 CTTTGTCAATTAATGCTCATATTTGGTGGGCACTTTAAATGAAATCCCGGTGACAAAC 1139  
DB 1098 CTTTGTCAATTAATGCTCATATTTAGTGGGCACTTTAAATGAAATCCCGGTGACAAAC 1157  
QY 1140 GAGAGAAAGTGGGAGTACGCTCAAGTCTCTGATGAGCCCTTAATGGGTAGGGCTTACACGTA 1199  
DB 1158 GAGAGAAAGTGGGAGTACGCTCAAGTCTCTGATGAGCCCTTAATGGGTAGGGCTTACACGTA 1217  
QY 1200 ATACATATGGCGGTACAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAAGCGC 1359  
DB 1218 ATACATATGGCGGTACAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAAGCGC 1277  
QY 1260 GTGTAGTCCGATTCGAGTCTGCAACTGCACTTCGTGAAGTGGAAATCGCTAGTAATCG 1319  
DB 1278 GTGTAGTCCGATTCGAGTCTGCAACTGCACTTCGTGAAGTGGAAATCGCTAGTAATCG 1337  
QY 1320 CGGATCAGCATGTTCGGGTGAATATCGTTCCCGGCTTTTGAACACACCGCCCGTACACCA 1379  
DB 1338 CGGATCAGCATGTTCGGGTGAATATCGTTCCCGGCTTTTGAACACACCGCCCGTACACCA 1397  
QY 1380 TGGAGTGGGTTTCAACAGAAAGCAGTATCTAAACGTAAGAGAGGCGCTTGCACCGGTG 1439  
DB 1398 TGGAGTGGGTTTCAACAGAAAGCAGTATCTAAACGTAAGAGAGGCGCTTGCACCGGTG 1457  
QY 1440 AGATTCAATGACTGGGGTG 1457  
DB 1458 AGATTCAATGACTGGGGTG 1475  
RESULT 12  
CQ796927  
LOCUS CQ796927 1491 bp DNA linear PAT 19-Apr-2004  
DEFINITION Sequence 20 from Patent WO2004026772.  
ACCESSION CQ796927  
VERSION CQ796927.1 GI:46408553

KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE unclassified sequences.  
1  
AUTHORS Hovenc, T.A.  
TITLE Ammonia-oxidizing bacteria and methods of using and detecting  
JOURNAL Patent: WO 2004026772-A 20 01-APR-2004;  
FEATURES  
source Location/Qualifiers  
1.1491  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="N. Aestuarii-like AOB Bp16clones7 16S rDNA"

ORIGIN  
Query Match 95.4%; Score 1390; DB 6; Length 1491;  
Best Local Similarity 97.5%; Pred. No. 8.9e-15;  
Matches 1422; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 1 ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTGAACGCGAGACGCGATGCTTGCAT 60  
DB 18 ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTGAACGCGAGACGCGATGCTTGCAT 77  
QY 61 CTGGTGGCGAGTGGCGGAGCGGATGATTAATGCTTCGGAACGTAATCCGAAAGAGGGGGTGA 120  
DB 78 CTGGTGGCGAGTGGCGGAGCGGATGATTAATGCTTCGGAACGTAATCCGAAAGAGGGGGTGA 137  
QY 121 ACGCATCGAAAGATGTGCTAATACCGCATATACCTTAAGGAGAAAGACGAGGATCGAAA 180  
DB 138 ACGCATCGAAAGATGTGCTAATACCGCATATACCTTAAGGAGAAAGACGAGGATCGAAA 197  
QY 181 GACCTTGCCTTTTGGAGCGCGCGATGTCTGATTAGCTAGTGGGCTTAAAGGCTTAC 240  
DB 198 GACCTTGCCTTTTGGAGCGCGCGATGTCTGATTAGCTAGTGGGCTTAAAGGCTTAC 257  
QY 241 CAAGGCGACGATCAAGTATGTTGCTTGAGAGAGACGACGACGACGCTGGGATCGAGCAG 300  
DB 258 CAAGGCGACGATCAAGTATGTTGCTTGAGAGAGACGACGACGACGCTGGGATCGAGCAG 317  
QY 301 GCCCAACTCTTACCGGAGGAGCAGCAGTGGGGAAATTTTGGCAATGGGCGCAAGCCTGATC 360  
DB 318 GCCCAACTCTTACCGGAGGAGCAGCAGTGGGGAAATTTTGGCAATGGGCGCAAGCCTGATC 377  
QY 361 CAGCAATGCGCGCTGAGTGAAGAGGCGCTTGGGTTGTAAGCTTTTCACTGAGAGA 420  
DB 378 CAGCAATGCGCGCTGAGTGAAGAGGCGCTTGGGTTGTAAGCTTTTCACTGAGAGA 437  
QY 421 AAAAGTTACGGTAAATTAATTCGTAACCCATGACCGTATTCGACGAAGAAGACCGGCTAAC 480  
DB 438 AAAAGTTACGGTAAATTAATTCGTAACCCATGACCGTATTCGACGAAGAAGACCGGCTAAC 497  
QY 481 TAGGTGCGACGAGCGCGGTAATACGTAAGGTCGCAAGCGTTAATCGAATTAATCTGGGCGT 540  
DB 498 TAGGTGCGACGAGCGCGGTAATACGTAAGGTCGCAAGCGTTAATCGAATTAATCTGGGCGT 557  
QY 541 AAAAGGTGGCGAGCGCGCTTGTAGTCAAGTGAATTCGCCGGGCTTAACTGGGAAT 600  
DB 558 AAAAGGTGGCGAGCGCGCTTGTAGTCAAGTGAATTCGCCGGGCTTAACTGGGAAT 617  
QY 601 TGGCTTTGAAATTAATCAAAAGTGAAGTGGCAGAGGAGGTGAATTCATGTTGTCAG 660  
DB 618 TGGCTTTGAAATTAATCAAAAGTGAAGTGGCAGAGGAGGTGAATTCATGTTGTCAG 677  
QY 661 TGAATGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 720  
DB 678 TGAATGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 737  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAAGTATCCTGTGATGTCACGCG 780  
DB 738 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAAGTATCCTGTGATGTCACGCG 797

QY 781 CTTAAACGATGCACTAGTGTGGGCGCTTATTAGGCTTGTGTAAGCAAGCTAACCGGTGA 840  
DB 798 CTTAAACGATGCACTAGTGTGGGCGCTTATTAGGCTTGTGTAAGCAAGCTAACCGGTGA 857  
QY 841 AGTTGACCGCTCGGGAGAGTACCGGTGCAAGATTAATACTCAAGGAATGACGGGAGACC 900  
DB 858 AGTTGACCGCTCGGGAGAGTACCGGTGCAAGATTAATACTCAAGGAATGACGGGAGACC 917  
QY 901 GCAACAAGCGGTGATTAATGATTAATTCATGCAAGCGGAAAACTTACTACTCCCTT 960  
DB 918 GCACAAGCGGTGATTAATGATTAATTCATGCAAGCGGAAAACTTACTACTCCCTT 977  
QY 961 GACATGTAGCGCAATTTTCTAGAGATGATGATG-CTTGGGAGCGCTTAACAGAGTGT 1019  
DB 978 GACATGTAGCGCAATTTTCTAGAGATGATGATG-CTTGGGAGCGCTTAACAGAGTGT 1037  
QY 1020 GCATGCTGCTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079  
DB 1038 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1097  
QY 1080 CTTGTGCTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139  
DB 1098 CTTGTGCTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157  
QY 1140 GGAGGAAGGTGGGAGTGAAGTCAAGTCCCTCATGCGCTTATGAGGAGCTTCAACAGTGA 1199  
DB 1158 GGAGGAAGGTGGGAGTGAAGTCAAGTCCCTCATGCGCTTATGAGGAGCTTCAACAGTGA 1217  
QY 1200 ATACAATGCGCGCTGAAGAAGGCTTCCAAACCGCGAGGAGGAGCTTAATCTCAAGAAAGCGC 1259  
DB 1218 ATACAATGCGCGCTGAAGAAGGCTTCCAAACCGCGAGGAGGAGCTTAATCTCAAGAAAGCGC 1277  
QY 1260 GTCTGATGCTCGGATGCGAGTCTGCACTGCACTGCTGTAAGTGGGAATTCGTAATTCG 1319  
DB 1278 GTCTGATGCTCGGATGCGAGTCTGCACTGCACTGCTGTAAGTGGGAATTCGTAATTCG 1337  
QY 1320 CGGATCAGCAGTGGCGGAGTGAAGTCCCGGAGCTTGTGTAACACGCGCGCTGACACCA 1379  
DB 1338 CGGATCAGCAGTGGCGGAGTGAAGTCCCGGAGCTTGTGTAACACGCGCGCTGACACCA 1397  
QY 1380 TGGAGTGGGTTTCAACGAAGCAGTGTCTTAACCGTAAGGAGGCGCTTGGCAGCGTG 1439  
DB 1398 TGGAGTGGGTTTCAACGAAGCAGTGTCTTAACCGTAAGGAGGCGCTTGGCAGCGTG 1457  
QY 1440 AGATTCAATGACTGGGCTG 1457  
DB 1458 AGATTCAATGACTGGGCTG 1475

RESULT 13  
AJ621032  
LOCUS Nitrosomonas sp. Is343 16S rRNA gene, isolate Is343. 1445 bp DNA linear BCT 08-JUN-2004  
DEFINITION  
ACCESSION AJ621032.1 GI:4094851  
VERSION 16S ribosomal RNA: 16S rRNA gene.  
KEYWORDS Nitrosomonas sp. Is343  
SOURCE Nitrosomonas sp. Is343  
ORGANISM Nitrosomonadaceae; Betaproteobacteria; Nitrosomonadales;  
Bacteria; Proteobacteria; Nitrosomonas.

REFERENCE  
1 Bollmann, A., Schmidt, I. and Bodelier, P. L. B.  
TITLE Influence of salt on activity, growth, protein pattern and fatty acid composition of the new isolated ammonia oxidizing betain Nitrosomonas strain Is343  
JOURNAL Unpublished  
AUTHORS Bollmann, A.  
REFERENCE Direct Submission  
TITLE Submitted (15-JAN-2004) Bollmann A.; Department for Microbial Ecology, NIOO-KNAW Centre for Limnology, Rijksstraatweg 6, 3631 AC Nieuwersluis, NETHERLANDS

FEATURES		Location/Qualifiers	
source	1..145	/organism="Nitrosomonas sp. Is343"	
		/mol_type="genomic DNA"	
		/isolate="Is343"	
		/isolation_source="brackish water"	
		/db_xref="taxon:261298"	
gene	1..144	/country="Netherlands:River Schelde"	
rRNA	/gene="16S rRNA"		
	/gene="16S rRNA"		
	/product="16S ribosomal RNA"		
ORIGIN			
Query Match	95.1%;	Score 1385.8;	DB 1; Length 1445;
Best Local Similarity	98.3%;	Pred. No. 1.1e-14;	
Matches 1421;	Conservative 0;	Mismatches 23;	Indels 2; Gaps 2;
Qy	5	AACGCTGGCGGCATGCTTTTACATGTCAGTCGACCGGACGACCGATGCTTGATCTGG	64
Db	1	AACGCTGGCGGCATGCTTTTACATGTCAGTCGACCGGACGACCGGATGCTTGACCTGG	60
Qy	65	TGCGAGTGGCGGACGGGTGATATGTCATCGGAACGTATCCAGAAAGGGGGTTAACG	124
Db	61	TGCGAGTGGCGGACGGGTGATATATCATCGGAACGTATCCAGAAAGGGGGATTACGC	120
Qy	125	ATCGAAAGATGTCTATACGGCATATCTTAAGAGGAAAGCAGGGGATCCGAAGACC	184
Db	121	ATCGAAAGATGTCTATATACCGCATATCTTGAGAGGAAAGCAGGGGATCCGAAGACC	180
Qy	185	TTGCGCTTTTGAAGCGGCGCATGCTGATTTAGCTAGTTGGTGGGTAAAGGCTTACCAAG	244
Db	181	TTGCGCTTTTGAAGCGGCGCATGCTGATTTAGCTAGTTGGTGGGTAAAGGCTTACCAAG	240
Qy	245	GCGACGATCAGTATGTTGCTTGAGAGGACGACCGACCACTGGGACTGAGACCGGCC	304
Db	241	GCGACGATCAGTATGTTGCTTGAGAGGACGACCGACCACTGGGACTGAGACCGGCC	300
Qy	305	AGACTCTTACGGAGGAGGAGAGTGGGGAAATTTGGACAAATGGGGCGCAAGCTGATCCAG	364
Db	301	AGACTCTTACGGAGGAGGAGAGTGGGGAAATTTGGACAAATGGGGCGCAAGCTGATCCAG	360
Qy	365	AATGCCGCTGAGTGAAGAAAGGCTTCGGGTGTTAAAGCTCTTTCAGTCGAGAAAGAAAG	424
Db	361	AATGCCGCTGAGTGAAGAAAGGCTTCGGGTGTTAAAGCTCTTTCAGTCGAGAAAGAAAG	419
Qy	425	GTTACGGTAAATATTCGTGACCCATGACGGTATCGACAGAAAGCACCGGCTTAATACG	484
Db	420	GTTACGGTAAATATTCGTGATTCATGACGGTATCGACAGAAAGCACCGGCTTAATACG	479
Qy	485	TGCCAGACGCGCGGTATATCTGAGGTGCAAGGCTTAATCGGAATTAATCGGGGGTAAAG	544
Db	480	TGCCAGACGCGCGGTATATCTGAGGTGCAAGGCTTAATCGGAATTAATCGGGGGTAAAG	539
Qy	545	GGTGGCAGGCGGCTTGTATGATGATGTAATCCCGGGCTTAACTTGGGAATTTGCG	604
Db	540	GGTGGCAGGCGGCTTGTATGATGATGTAATCCCGGGCTTAACTTGGGAATTTGCG	599
Qy	605	TTTGAATCTACAAAGCTAGAGTGTGGCAGAGGAGTGAATTCATGTGTAGCAGTGA	664
Db	600	TTTGAATCTACAAAGCTAGAGTGTGGCAGAGGAGTGAATTCATGTGTAGCAGTGA	659
Qy	665	ATGCGTAGATATGGAAGAACTTCGATGGGAGGCGGCTCCGGGTTAACTGACG	724
Db	660	ATGCGTAGATATGGAAGAACTTCGATGGGAGGCGGCTCCGGGTTAACTGACG	719
Qy	725	CTCATGACGAAAGGCTGGGAGCAACAGATTAGATACCTGGTATCCACGCCCTAA	784
Db	720	CTCATGACGAAAGGCTGGGAGCAACAGATTAGATACCTGGTATCCACGCCCTAA	779
Qy	785	ACGATGTCAACTAGTTGTTGGGCTTATTAAGCTTGTGAAGAACTAACCGGTGAAGTT	844

Db	780	ACGATGTCAACTAGTTGTTGGGCTTATTAAGCTTGTGAAGAACTAACCGGTGAAGTT	839
Qy	845	GAACCGCTTGGGAGTACGGTCCGAAAGATTAAATCTAAAGAAATTGACGGGGACCGGCAC	904
Db	840	GAACCGCTTGGGAGTACGGTCCGAAAGATTAAATCTAAAGAAATTGACGGGGACCGGCAC	899
Qy	905	AAGCGGTGATTTATGATTAATTCATGTCAGAACGGAAGAAACCTTACCTACCTTTGACA	964
Db	900	AAGCGGTGATTTATGATTAATTCATGTCAGAACGGAAGAAACCTTACCTACCTTTGACA	959
Qy	965	TGTAGCGAATTTCTAGAGATAGATTAGTG-CTTCGGAAACGTTAACAGAGTGTGCAT	1023
Db	960	TGTAGCGAATTTCTAGAGATAGATTAGTGCTTCGGAAACGTTAACAGAGTGTGCAT	1019
Qy	1024	GCGTGTGTCAGCTCGTGTCTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAACCTT	1083
Db	1020	GCGTGTGTCAGCTCGTGTCTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAACCTT	1079
Qy	1084	GTCATTAATTTGCCATCATTTGGTGGGCACTTTAATGAGACTGCCGGTGCAGAACCGGAG	1143
Db	1080	GTCATTAATTTGCCATCATTTAGTTGGGCACTTTAATGAGACTGCCGGTGCAGAACCGGAG	1139
Qy	1144	GAAGTGGGAGTACGTCAGATCCCTCATGAGCCCTTAATGGGTAGGCTTCAACAGTAATAC	1203
Db	1140	GAAGTGGGAGTACGTCAGATCCCTCATGAGCCCTTAATGGGTAGGCTTCAACAGTAATAC	1199
Qy	1204	AATGGCGCTGACAGAGGTTGCCAACCCCGAGGGGAGTAACTTCAGAAACCGGCTG	1263
Db	1200	AATGGCGCTGACAGAGGTTGCCAACCCCGAGGGGAGTAACTTCAGAAACCGGCTG	1259
Qy	1264	TATCCCGGATCCGAGTCTGCACTCGAATCCGGAATGCGGAATGCTAATTCGCGGA	1323
Db	1260	TATCCCGGATCCGAGTCTGCACTCGAATCCGGAATGCGGAATGCTAATTCGCGGA	1319
Qy	1324	TCAGCATGTGCGCGGTAAATAGTTCGCGGCTTGTATACACACCGCGCTCACACATGGG	1383
Db	1320	TCAGCATGTGCGCGGTAAATAGTTCGCGGCTTGTATACACACCGCGCTCACACATGGG	1379
Qy	1384	AGTGGGTTTACAGAAAGCAGTATGCTTAACCTTAAGAGAGGGGCTTGCACCGGTGAT	1443
Db	1380	AGTGGGTTTACAGAAAGCAGTATGCTTAACCTTAAGAGAGGGGCTTGCACCGGTGAT	1439
Qy	1444	TCATGA 1449	
Db	1440	TCATGA 1445	
RESULT 14			
AF272420			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
Location/Qualifiers			
1450 bp DNA linear BCT 05-DEC-2000			
Nitrosomonas aestuarii 16S ribosomal RNA gene, partial sequence.			
AF272420			
AF272420.1 GI:11545280			
Nitrosomonas aestuarii			
Nitrosomonas aestuarii			
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;			
Nitrosomonadaceae; Nitrosomonas.			
1 (bases 1 to 1450)			
Purkhold, U., Pommerening-Rozer, A., Juretschko, S., Schmid, M. C.,			
Koops, H. P., and Wagner, M.			
Phylogeny of all recognized species of ammonia oxidizers based on			
comparative 16S rRNA and amoA sequence analysis: implications for			
molecular diversity surveys			
Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)			
11097916			
2 (bases 1 to 1450)			
Purkhold, U., Pommerening-Rozer, A., Juretschko, S., Schmid, M. C.,			
Koops, H. P., and Wagner, M.			
Direct Submission			
Submitted (26-MAY-2000) Lehrstuhl fuer Mikrobiologie, Technische			
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany			

```
source
1..1450
/organism="Nitrosomonas aestuarii"
/mol_type="genomic DNA"
/db_xref="taxon:52441"
<1..>1450
/product="16S ribosomal RNA"
ORIGIN
Query Match 94.9%; Score 1382; DB 1; Length 1450;
Best Local Similarity 97.5%; Pred. No. 1,2e-14;
Matches 1414; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
QY 2 TTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAAACGAGACGAGATGCTTCATC 61
DB 1 TTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAAACGAGACGAGATGCTTCATC 60
QY 62 TGTGCGCAGTGGCGGACGCGGTGATATGATCGAACGTATCCAGAAAGGGGGGTA 121
DB 61 TGTGCGCAGTGGCGGACGCGGTGATATGATCGAACGTATCCAGAAAGGGGGGTA 120
QY 122 CGCATGAAAGATGTCTTAATACCGCATATCTTAAGAGAGAAAGCAGGATCGAAAG 181
DB 121 CGCATGAAAGATGTCTTAATACCGCATATCTTAAGAGAGAAAGCAGGATCGAAAG 180
QY 182 ACCTTGCGCTTTGAGCGCGCGATGCTGATTAAGCTAGTGGTGGGGTAAAGGCTTACC 241
DB 181 ACCTTGCGCTTTGAGCGCGCGATGCTGATTAAGCTAGTGGTGGGGTAAAGGCTTACC 240
QY 242 AAGGCAGCATCATGATGTTGCTGAGAGAGACGACGACCACTGGGACTGAGACAGG 301
DB 241 AAGGCAGCATCATGATGTTGCTGAGAGAGACGACGACCACTGGGACTGAGACAGG 300
QY 302 CCCAGCTCTTACGGGAGCAGACGATGGGGAATTTTGAACAATGGGCGCAAGCTTATCC 361
DB 301 CCCAGCTCTTACGGGAGCAGACGATGGGGAATTTTGAACAATGGGCGCAAGCTTATCC 360
QY 362 AGCAATGCGCGCTGATGAGAAAGGCTTCGCGTGTAAAGCTCTTCACTGAGAGAA 421
DB 361 AGCAATGCGCGCTGATGAGAAAGGCTTCGCGTGTAAAGCTCTTCACTGAGAGAA 420
QY 422 AAGGTTACGGTAATATATCGTGAACCATGACGGTATCGACAGAAAGAACCGGCTAACT 481
DB 421 AAGGTTACGGTAATATATCGTGAACCATGACGGTATCGACAGAAAGAACCGGCTAACT 480
QY 482 AGGTGCGACAGACCGCGGTATATACGTAGGGTGAACGCTTAATCGAAATTAATCGGGCGTA 541
DB 481 AGGTGCGACAGACCGCGGTATATACGTAGGGTGAACGCTTAATCGAAATTAATCGGGCGTA 540
QY 542 AAGGTTGCGCAGCGGCTTGTAAAGTCAGATGGAATCCCGGGCTTAACTGGGAAT 601
DB 541 AAGGTTGCGCAGCGGCTTGTAAAGTCAGATGGAATCCCGGGCTTAACTGGGAAT 600
QY 602 GCGTTGAAACTACAAAGCTAGAGTGGCAGAGGAGGTGGAATTCATGTGTAGCACT 661
DB 601 GCGTTGAAACTACAAAGCTAGAGTGGCAGAGGAGGTGGAATTCATGTGTAGCACT 660
QY 662 GAAATGCGTGAAGTATGGAAGAACTCATGTGCGGAAGCAGCTCTCTGGTTAACTACACTG 721
DB 661 GAAATGCGTGAAGTATGGAAGAACTCATGTGCGGAAGCAGCTCTCTGGTTAACTACACTG 720
QY 722 AGGCTCATSCAGAAAGCGTGGGAGCAACAGATTAATACCTCGTGTAGTCAGCGCC 781
DB 721 AGGCTCATSCAGAAAGCGTGGGAGCAACAGATTAATACCTCGTGTAGTCAGCGCC 780
QY 782 TAAACGATGTCAATGATTGTTGGGCTTATTAGGCTTGTATACGAGCTAAACGCGTGA 841
DB 781 TAAACGATGTCAATGATTGTTGGGCTTATTAGGCTTGTATACGAGCTAAACGCGTGA 840
QY 842 GTTGCACCGCTGGGAGTACGCTCGCAATTAATACTAAAGGAATTTGACGGGACCGG 901
DB 841 GTTGCACCGCTGGGAGTACGCTCGCAATTAATACTAAAGGAATTTGACGGGACCGG 900
QY 902 CACAAGCGGTGATTATGTGATTAATTCGATGCAACGCGAAAAACCTTACCTACCTTG 961
```

```
DB 901 CACAAGCGGTGATTATGTGATTAATTCGATGCAACGCGAAAAACCTTACCTACCTTG 960
QY 962 ACATGTAGCGAATTTTCTAGAGATGATTAGTG-CTTCGGGAACGTTAACACAGGTCTG 1020
DB 961 ACATGTAGCGAATTTTCTAGAGATGATTAGTG-CTTCGGGAACGTTAACACAGGTCTG 1020
QY 1021 CATGGCTGTCAGTCTGTCGTGAGATGTTGGTTAGTCCGCAACGAGCGCAACC 1080
DB 1021 CATGGCTGTCAGTCTGTCGTGAGATGTTGGTTAGTCCGCAACGAGCGCAACC 1080
QY 1081 CTGTGATTAATTTGCAATCATTTGGTGGGCACTTTAATGACACTCCGGTGAACAAACCG 1140
DB 1081 CTGTGATTAATTTGCAATCATTTAGTGGGCACTTTAATGACACTCCGGTGAACAAACCG 1140
QY 1141 GAGGAAGGTGGGATGACGCTCAAGTCTCATGSCCTTATGGGTAGAGCTTACACGTAA 1200
DB 1141 GAGGAAGGTGGGATGACGCTCAAGTCTCATGSCCTTATGGGTAGAGCTTACACGTAA 1200
QY 1201 TACAATGGCGCTTACAGAGGCTTGCACACCGCGAGGGGAGCTAATCTCAAAAGCGCG 1260
DB 1201 TACAATGGCGCTTACAGAGGCTTGCACACCGCGAGGGGAGCTAATCTCAAAAGCGCG 1260
QY 1261 TCGTAGTCGCGATCGGAGTCTGCAACTGCACTCCGTGAAGTGGAAATGCTAGTAATCGC 1320
DB 1261 TCGTAGTCGCGATCGGAGTCTGCAACTGCACTCCGTGAAGTGGAAATGCTAGTAATCGC 1320
QY 1321 GGATCAGCATGTGCGGTGATATACCTTCGCGGCTTGTACACACCGCCCGTACACCAT 1380
DB 1321 GGATCAGCATGTGCGGTGATATACCTTCGCGGCTTGTACACACCGCCCGTACACCAT 1380
QY 1381 GGGAGTGGGTTTACACAGAGCAGTACTTAACCGTAAGAGGGCGCTTGGCACGCTGA 1440
DB 1381 GGGAGTGGGTTTACACAGAGCAGTACTTAACCGTAAGAGGGCGCTTGGCACGCTGA 1440
QY 1441 GATTCATGAC 1450
DB 1441 GATTCATGAC 1450
RESULT 15
AF272424 1513 bp DNA linear BCT 05-DEC-2000
LOCUS Nitrosomonas sp. NM 51 16S ribosomal RNA gene, partial sequence.
DEFINITION AF272424.1
ACCESSION AF272424.1 GI:11545284
VERSION AF272424.1
KEYWORDS
SOURCE
ORGANISM Nitrosomonas sp. Nms1
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosomonas.
REFERENCE
1 (bases 1 to 1513)
Purkhold,U., Pommerening-Roser,A., Juretschko,S., Schmid,M.C.,
Koops,H.P. and Wagner,M.
Phylogeny of all recognized species of ammonia oxidizers based on
comparative 16S rRNA and amoA sequence analysis: implications for
molecular diversity surveys
Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)
JOURNAL
PUBMED 11097916
REFERENCE
2 (bases 1 to 1513)
Purkhold,U., Pommerening-Roser,A., Juretschko,S., Schmid,M.C.,
Koops,H.P. and Wagner,M.
Direct Submission
Submitted (26-MAY-2000) Lehrstuhl fuer Mikrobiologie, Technische
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES
source
1..1513
/organism="Nitrosomonas sp. Nms1"
/mol_type="genomic DNA"
/db_xref="taxon:133720"
<1..>1513
/product="16S ribosomal RNA"
ORIGIN
```

Query Match 94.8%; Score 1380.6; DB 1; Length 1513;  
 Best Local Similarity 97.1%; Pred. No. 1.2e-14;  
 Matches 1415; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

```

QY 1 ATTGAACGCTGGCGGCAATGCTTTTACACATGCAAGTGGAAAGGCAACGCAATGCTTGGAT 60
   |||
DB 21 ATTGAACGCTGGCGGCAATGCTTTTACACATGCAAGTGGAAAGGCAACGCAATGCTTGGAT 80
   |||

QY 61 CTGGTGGCGAGTGGCGGCAATGCTTTTACACATGCAAGTGGAAAGGCAACGCAATGCTTGGAT 120
   |||
DB 81 CTGGTGGCGAGTGGCGGCAATGCTTTTACACATGCAAGTGGAAAGGCAACGCAATGCTTGGAT 140
   |||

QY 121 AGCATCGAAAGATGTGCTAATACCGCATATCTTAAAGAGGAAAGCAGGGGATCGAA 180
   |||
DB 141 AGCATCGAAAGATGTGCTAATACCGCATATCTTAAAGAGGAAAGCAGGGGATCGTA 200
   |||

QY 181 GACCTTGGCTTTTGAAGCGGCGGATGCTGATTTAGCTAGTGGTGGGTTAAAGGCTTAC 240
   |||
DB 201 GACCTTGGCTTTTGAAGCGGCGGATGCTGATTTAGCTAGTGGTGGGTTAAAGGCTTAC 260
   |||

QY 241 CAAGGCGAGATAGATAGTGTGCTGAGAGAGACGACGACCTGGGACGCTGAGACAG 300
   |||
DB 261 CAAGGCTACGATAGATAGTGTGCTGAGAGAGACGACGACCTGGGACGCTGAGACAG 320
   |||

QY 301 GCCCAGACTCTTACGGGAGGACAGAGTGGGAAATTTTGAACAATGGCGGCAAGCTTATC 360
   |||
DB 321 GCCCAGACTCTTACGGGAGGACAGAGTGGGAAATTTTGAACAATGGCGGCAAGCTTATC 380
   |||

QY 361 CAGCAATGCCGCTGATGAGAAAGGCTTTCGGGTTGAAAGCTTTTCACTGAGAGAA 420
   |||
DB 381 CAGCAATGCCGCTGATGAGAAAGGCTTTCGGGTTGAAAGCTTTTCACTGAGAGAA 440
   |||

QY 421 AAAGGTACGGTAAATATATGTCGATCCATGACGGTATCGACAGAAAGACCGGCTTAC 480
   |||
DB 441 AAAGATTTGATGATATATGTCGATCCATGACGGTATCGACAGAAAGACCGGCTTAC 500
   |||

QY 481 TAGCTGCACAGCGCGGCTAATCTAGAGGTGCAACGCTTAACTCGAATTTACTGGGCGT 540
   |||
DB 501 TAGCTGCACAGCGCGGCTAATCTAGAGGTGCAACGCTTAACTCGAATTTACTGGGCGT 560
   |||

QY 541 AAAGGTGCGCAGCGGCTTGTAAAGTCAATGTGAAATCCCGGCTTAACTCGGGAAT 600
   |||
DB 561 AAAGGTGCGCAGCGGCTTGTAAAGTCAATGTGAAATCCCGGCTTAACTCGGGAAT 620
   |||

QY 601 TGGCTTGAATCTAACAAGCTAGGTGCGAGAGGAGGAGTGAATTCATGTGTAGCAG 660
   |||
DB 621 TGGCTTGAATCTAACAAGCTAGGTGCGAGAGGAGGAGTGAATTCATGTGTAGCAG 680
   |||

QY 661 TGAATGCGTATGAGATATGGAAGCAATCGATGCGAAGGCAAGGCTCTCGGCTTAACT 720
   |||
DB 681 TGAATGCGTATGAGATATGGAAGCAATCGATGCGAAGGCAAGGCTCTCGGCTTAACT 740
   |||

QY 721 GACGCTATGACAGAAAGCGTGGGAGCAAAACAGATTTAGTACCTGATGTCACGCC 780
   |||
DB 741 GACGCTATGACAGAAAGCGTGGGAGCAAAACAGATTTAGTACCTGATGTCACGCC 800
   |||

QY 781 CTAAACGATGCACTAGTGTGTGGGCTTATTTAGGCTTGTGTAACGAAGCTTAAGCGCGTA 840
   |||
DB 801 CTAAACGATGCACTAGTGTGTGGGCTTATTTAGGCTTGTGTAACGAAGCTTAAGCGCGTA 860
   |||

QY 841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAAACTCAAGGAATTTGACGGGAGCC 900
   |||
DB 861 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAAACTCAAGGAATTTGACGGGAGCC 920
   |||

QY 901 GCACAAGCGGTGATTATGTGATTTAATTCATGCAACGCGAATACTTAACCTTACCTT 960
   |||
DB 921 GCACAAGCGGTGATTATGTGATTTAATTCATGCAACGCGAATACTTAACCTTACCTT 980
   |||

QY 961 GACATGTATGCAATTTTCTAGAGATATGTTAGTCTTGGGAAACGCTTAACACAGGTCT 1019
   |||
DB 981 GACATGTATGCAATTTTCTAGAGATATGTTAGTCTTGGGAAACGCTTAACACAGGTCT 1040
   |||

```

```

QY 1020 GGATGGCTGTGCTGACGCTGTGTCGAGATGTTGGGTTAAGTCCCGACAGCGCCAAC 1079
   |||
DB 1041 GGATGGCTGTGCTGACGCTGTGTCGAGATGTTGGGTTAAGTCCCGACAGCGCCAAC 1100
   |||

QY 1080 CCTTGCATTTAATTTGCAATCTTTGTTGGGCACTTTAATGAGACTCCGGTGACAAACC 1139
   |||
DB 1101 CCTTGCATTTAATTTGCAATCTTTGTTGGGCACTTTAATGAGACTCCGGTGACAAACC 1160
   |||

QY 1140 GGAGAGAGTGGGAGTACGCTGAACTCTCAATGACCCCTTAATGAGTGGGCTTACACGTA 1199
   |||
DB 1161 GGAGAGAGTGGGAGTACGCTGAACTCTCAATGACCCCTTAATGAGTGGGCTTACACGTA 1220
   |||

QY 1200 ATTCATGTCGCTGATACAGGGGTTCACACCGCGAGGGGAGCTAATCTCAGAAAGCCG 1259
   |||
DB 1221 ATTCATGTCGCTGATACAGGGGTTCACACCGCGAGGGGAGCTAATCTCAGAAAGCCG 1280
   |||

QY 1260 GTCTAGTCCCGATCGGAGTCTGCACTCGACTCCGTGAAGTGGAAATCGTAAATCG 1319
   |||
DB 1281 GTCTAGTCCCGATCGGAGTCTGCACTCGACTCCGTGAAGTGGAAATCGTAAATCG 1340
   |||

QY 1320 CGGATCAGATGTGCGGTGAATACGTTCCCGGCTTTGTACACACCCCGCTCAACCA 1379
   |||
DB 1341 CGGATCAGATGTGCGGTGAATACGTTCCCGGCTTTGTACACACCCCGCTCAACCA 1400
   |||

QY 1380 TGGAGTGGGTTTACACAGAGAGGATGTTAACGTTAAGAGAGGGGCTTGGCACGGTG 1439
   |||
DB 1401 TGGAGTGGGTTTACACAGAGAGGATGTTAACGTTAAGAGAGGGGCTTGGCACGGTG 1460
   |||

QY 1440 AGATTCATGACTGGGGTG 1457
   |||
DB 1461 AGATTCATGACNNNGTG 1478
   |||

```

RESULT 16  
 AF272418 1492 bp DNA linear BCT 05-DEC-2000  
 LOCUS  
 DEFINITION Nitrosomonas marina 16S ribosomal RNA gene, partial sequence.  
 ACCESSION AF272418  
 VERSION AF272418.1 GI:11545278  
 KEYWORDS  
 SOURCE Nitrosomonas marina  
 ORGANISM Nitrosomonas marina  
 Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
 Nitrosomonadaceae; Nitrosomonas.

REFERENCE  
 AUTHORS Purkhold,U., Pommerening-Roser,A., Juretschko,S., Schmid,M.C.,  
 1 (bases 1 to 1492)  
 TITLE Phylogeny of all recognized species of ammonia oxidizers based on  
 comparative 16S rRNA and amoA sequence analysis: implications for  
 molecular diversity surveys  
 JOURNAL Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)  
 PUBMED 11097916  
 REFERENCE 2 (bases 1 to 1492)  
 AUTHORS Purkhold,U., Pommerening-Roser,A., Juretschko,S., Schmid,M.C.,  
 Koops,H.P. and Wagner,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-MAY-2000) Leibniz-Institut fuer Mikrobiologie, Technische  
 Universitaet Muenchen, Am Hoechinger 4, Freising 85350, Germany

FEATURES  
 source  
 1..1492  
 /organism="Nitrosomonas marina"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:917"  
 <1..>1492  
 /product="16S ribosomal RNA"

ORIGIN  
 Query Match 94.8%; Score 1378.6; DB 1; Length 1492;  
 Best Local Similarity 97.2%; Pred. No. 1.3e-14;  
 Matches 1413; Conservative 0; Mismatches 39; Indels 1; Gaps 1;  
 6 ACGTGGCGGATGCTTTTACATGCAAGTGGAAAGGCAACGCAACGATGCTTGCATCTGT 65  
 |||

Db 1 AGCGTGGCGGACGCTTTACATGCAAGTGCAGACGGGAGCGGCTTGCACTCGGT 60  
Qy 66 GGGGAGTGGCGGACGGGTGAGTATGATCGGAACGTATTCAGAAAGAGGGGGGTAAACGA 125  
Db 61 GGGGAGTGGCGGACGGGTGAGTATGATCGGAACGTATTCAGAAAGAGGGGGGTAAACGA 120  
Qy 126 TCGAAAGATGTGTAAATACCGCATATCTAAGAGAGAAAGCAGGGGATCGAAAGACCT 185  
Db 121 TCGAAAGATGTGTAAATACCGCATATCTAAGAGAGAAAGCAGGGGATCGTAACACT 180  
Qy 186 TGGGCTTTTGGAGTGGCCGCGATCTGATTAGCTTAGTGGGTAAAGGCTTACCAAG 245  
Db 181 TGGGCTTTTGGAGTGGCCGCGATCTGATTAGCTTAGTGGGTAAAGGCTTACCAAG 240  
Qy 246 CGAGCATCAGTAGTGTGTCTGAGAGAGCAACGACCACTGGGACTGAGACACGGCCCA 305  
Db 241 CGAGCATCAGTAGTGTGTCTGAGAGAGCAACGACCACTGGGACTTAAAGACGGCCCA 300  
Qy 306 GACTCTTACGGGAGGAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCCTGATCCAGCA 365  
Db 301 GACTCTTACGGGAGGAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCCTGATCCAGCA 360  
Qy 366 ATGCCCGCTGAGTGAAGAAAGGCTTCGGGTGTAAAGCTTTTCAGTCGAGAAAGAAAG 425  
Db 361 ATGCCCGCTGAGTGAAGAAAGGCTTCGGGTGTAAAGCTTTTCAGTCGAGAAAGAAAG 420  
Qy 426 TTAAGCTTAATATTCGTGACCCATGACGCTATTCGACAGAAAGAACCGGCTTAACCT 485  
Db 421 TTAAGCTTAATATTCGTGACCCATGACGCTATTCGACAGAAAGAACCGGCTTAACCT 480  
Qy 486 GCAGAGCGCGCGGTATATCGTAGGTGCAAGGCTTAATCGAAATTAAGGCGGTAAAG 545  
Db 481 GCAGAGCGCGCGGTATATCGTAGGTGCAAGGCTTAATCGAAATTAAGGCGGTAAAG 540  
Qy 546 GTGCGGAGCGGCTTTGTAAGTCAATGTAATCCCGGCTTAACTGGGAATTCGT 605  
Db 541 GTGCGGAGCGGCTTTGTAAGTCAATGTAATCCCGGCTTAACTGGGAATTCGT 600  
Qy 606 TTGAAACTTCAAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAGTGA 665  
Db 601 TTGAAACTTCAAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAGTGA 660  
Qy 666 TCGGTGAGATATGGAAGAACATCGATGGGAAAGGAGGCTTCGGGTAAACCTGACG 725  
Db 661 TCGGTGAGATATGGAAGAACATCGATGGGAAAGGAGGCTTCGGGTAAACCTGACG 720  
Qy 726 TCATGACGAAAGCGTGGGAGCAACAGATTTAGTACCTGTGATGTCACGCGCTTAA 785  
Db 721 TCATGACGAAAGCGTGGGAGCAACAGATTTAGTACCTGTGATGTCACGCGCTTAA 780  
Qy 786 CGATGTCACCTAGTGTGTGGGCTTATTAAGGCTTGTGTAACGAAGCTTAACGCGTGA 845  
Db 781 CGATGTCACCTAGTGTGTGGGCTTATTAAGGCTTGTGTAACGAAGCTTGAAGGTTG 840  
Qy 846 ACCGCTGGGAGTACGCTCGCAAGATTTAAACTCAAGGAATTTGACGGGAGCCGCA 905  
Db 841 ACCGCTGGGAGTACGCTCGCAAGATTTAAACTCAAGGAATTTGACGGGAGCCGCA 900  
Qy 906 AGCGGTGATTAATGTGATTAATTCGATCAACGCGAAAAACCTTAACCTTACCTTGA 965  
Db 901 AGCGGTGATTAATGTGATTAATTCGATCAACGCGAAAAACCTTAACCTTACCTTGA 960  
Qy 966 GTAGGGAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTAACACAGGTCTG 1024  
Db 961 GTAGGGAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTAACACAGGTCTG 1020  
Qy 1025 GCTGTGTCAGCTCGTGTGAGTGAAGTGGGTTAAGCCCGCAACGAGCGCAACCTTG 1084  
Db 1021 GCTGTGTCAGCTCGTGTGAGTGAAGTGGGTTAAGCCCGCAACGAGCGCAACCTTG 1080  
Qy 1085 TCATTAATTCGATCAATTTGTTGGCACTTTAATGAGCTGCGGTTGACAAACGGGAG 1144  
Db 1081 TCATTAATTCGATCAATTTGTTGGCACTTTAATGAGCTGCGGTTGACAAACGGGAG 1140

Qy 1145 AAGTGGGAGATGACGTCAAGTCTCTATGCGCCCTTAATGGGTAGGGCTTCAACAGTAATCA 1204  
Db 1141 AAGTGGGAGATGACGTCAAGTCTCTATGCGCCCTTAATGGGTAGGGCTTCAACAGTAATCA 1200  
Qy 1205 ATGGCGGTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAACGGCTGT 1264  
Db 1201 ATGGCGGTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAACGGCTGT 1260  
Qy 1265 AGTCCGATCGGAGTCTGCAATCTGCACTCCGTAAGTGGGAATCGCTAGTAATCCGGAT 1324  
Db 1261 AGTCCGATCGGAGTCTGCAATCTGCACTCCGTAAGTGGGAATCGCTAGTAATCCGGAT 1320  
Qy 1325 CAGCATGTGCGGGTGAATTCGTTCCGGGCTTGTACACACCGCCCGTCACACATGGGA 1384  
Db 1321 CAGCATGTGCGGGTGAATTCGTTCCGGGCTTGTGTACACACCGCCCGTCACACATGGGA 1380  
Qy 1385 GTGGGTTTCAACGAGAGAGTATCTAACCTGAAGAGGGGCTTGGCACCAGTGAAT 1444  
Db 1381 GTGGGTTTCAACGAGAGAGTATCTAACCTGAAGAGGGGCTTGGCACCAGTGAAT 1440  
Qy 1445 CATGACTGGGGTG 1457  
Db 1441 CATGACTGGGGTG 1453

RESULT 17  
CQ796926 1494 bp DNA linear PAT 19-APR-2004

LOCUS CQ796926 Sequence 19 from Patent WO2004026772.  
DEFINITION CQ796926  
ACCESSION CQ796926  
VERSION CQ796926.1 GI:46408552  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 unidentified sequences.  
AUTHORS Hovanec, T. A.  
TITLE Ammonia-oxidizing bacteria and methods of using and detecting  
JOURNAL Patent: WO 2004026772-A 19 01-APR-2004;  
Aquadria Inc. (US)  
FEATURES  
source 1. 1494  
Location/Qualifiers  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="N. aestuarii-like AOB Paclomej1 16S rDNA"

ORIGIN

Query Match 93.2%; Score 1357.6; DB 6; Length 1494;  
Best Local Similarity 96.2%; Pred. No. 2.8e-14;  
Matches 1401; Conservative 0; Mismatches 54; Indels 1; Gaps 1;

Qy 1 ATTTGACGCTGGCGGATGCTTTTACATGCAATGCAAGTCGAAACGACGAGATGCTTGAT 60  
Db 19 ATTTGACGCTGGCGGATGCTTTTACATGCAATGCAAGTCGAAACGACGAGATGCTTGAT 78  
Qy 61 CTGTGGCGAGTGGCGGACGGGTGATATGATCGAAACGTAATCCAGAGAGGGGGTTA 120  
Db 79 CTGTGGCGAGTGGCGGACGGGTGATATGATCGAAACGTAATCCAGAGAGGGGGTTA 138  
Qy 121 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGAGGGGATCGAAA 180  
Db 139 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGAGGGGATCGAAA 198  
Qy 181 GACCTTGGCTTTTGGAGCGGCGCGATGTCTGATTAGTAGTGTGGGTAAAGGCTTAC 240  
Db 199 GACCTTGGCTTTTGGAGCGGCGCGATGTCTGATTAGTAGTGTGGGTAAAGGCTTAC 258  
Qy 241 CAAGCGACGATCAAGTATGTTGGTTGAGAGAGCAACGCAACTGGGACTGAGACACG 300  
Db 259 CAAGCGACGATCAAGTATGTTGGTTGAGAGAGCAACGCAACTGGGACTGAGACACG 318

QY 301 GCCCAGACTCTTACGGGAGCAGAGTGGGAAATTTTGCACATGGGCGCAACCTGATC 360  
 DB 319 GCCCAGACTCTTACGGGAGCAGAGTGGGAAATTTTGCACACGGGCGAAACCTGATC 378  
 QY 361 CAGCAATGCGCGGTAGTGAAGAAGGCTTGGGTTTGAAGCTCTTTCACTGAGAGA 420  
 DB 379 CAGCAATGCGCGGTAGTGAAGAAGGCTTGGGTTTGAAGCTCTTTCACTGAGAGA 438  
 QY 421 AAAAGTTACGGTAAATATGTGACCATGACGGTATCGACAGAGAAGCAGCGCTAAC 480  
 DB 439 AAAAGTTGATTAATATCAACTTATGACGTTACCGACAGAGAAGCAGCGCTAAC 498  
 QY 481 TACGTCCAGCAGCCCGGTATACGTAAGGTGCAAGCGTTAAATCGAATTTACTGGCGT 540  
 DB 499 TACGTCCAGCAGCCCGGTATACGTAAGGTGCAAGCGTTAAATCGAATTTACTGGCGT 558  
 QY 541 AAAAGGTGCGCAGCGCGCTTGTAAAGTCAAGTGAATCCCGGCTTAACTGGGAAT 600  
 DB 559 AAAAGGTGCGCAGCGCGCTTGTAAAGTCAAGTGAATCCCGGCTTAACTGGGAAT 618  
 QY 601 TGGGTTGAAACTCAAAAGCTAGAGTGTGCAAGGAGGTGAATTCATGTGTAGCAG 660  
 DB 619 TGGGTTGAAACTCAAAAGCTAGAGTGTGCAAGGAGGTGAATTCATGTGTAGCAG 678  
 QY 661 TGAATGCGGTAGATATGGAAGACATCGATGGCGAAGCAGCGCTTCTGGGTTAACT 720  
 DB 679 TGAATGCGGTAGATATGGAAGACATCGATGGCGAAGCAGCGCTTCTGGGTTAACT 738  
 QY 721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATCCTGTGTATGTCACGCC 780  
 DB 739 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATCCTGTGTATGTCACGCC 798  
 QY 781 CTAAACGATGCACTAGTGTGTGGGCTTATTAAGCTTGTAAAGCAAGCTAAAGCTGA 840  
 DB 799 CTAAACGATGCACTAGTGTGTGGGCTTATTAAGCTTGTAAAGCTAAAGCTGA 858  
 QY 841 AGTTGACCGCTGGGAGTACGGTGGCAAGATTAAACTCAAGAAATTGACGGGAGCC 900  
 DB 859 AGTTGACCGCTGGGAGTACGGTGGCAAGATTAAACTCAAGAAATTGACGGGAGCC 918  
 QY 901 GCAACAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAATACTTAACCTT 960  
 DB 919 GCAACAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAATACTTAACCTT 978  
 QY 961 GACATGAGCAATTTTCTAGAGATAGATTAGT-CTTCGGGAGCGCTAACAGAGTGT 1019  
 DB 979 GACATGAGCAATTTTCTAGAGATAGATTAGT-CTTCGGGAGCGCTAACAGAGTGT 1038  
 QY 1020 GCATGGCTGTCTGAGCTGTGTGTGATGTTGGGTTAAATCCCGCAACGAGCGCAAC 1079  
 DB 1039 GCATGGCTGTCTGAGCTGTGTGTGATGTTGGGTTAAATCCCGCAACGAGCGCAAC 1098  
 QY 1080 CTTTGTATTAAATTTGCAATCTTTGTTGGGCACTTTAATGAGACTGCGGTTGCAAAAC 1139  
 DB 1099 CTTTGTATTAAATTTGCAATCTTTAATGAGACTGCGGTTGCAAAAC 1158  
 QY 1140 GAGAGAAAGTGGGATGACGTAAGTCTCTATGAGCTTAAAGGCTTTCACAGCTA 1199  
 DB 1159 GAGAGAAAGTGGGATGACGTAAGTCTCTATGAGCTTAAAGGCTTTCACAGCTA 1218  
 QY 1200 ATACAAATGGCGGTACAGAGGTTGCCAAACCGCGAGGGGAGCTTAATCTGAGAAAGGC 1259  
 DB 1219 ATACAAATGGCGGTACAGAGGTTGCCAAACCGCGAGGGGAGCTTAATCTGAGAAAGGC 1278  
 QY 1260 GTCTGATGTCGGATGCAAGTCTGCAACTGACTCGGTGAAGTGGAAATGCTAGTAATCG 1319  
 DB 1279 GTCTGATGTCGGATGCAAGTCTGCAACTGACTCGGTGAAGTGGAAATGCTAGTAATCG 1338  
 QY 1320 CGGATCAGCAATGTCCGGGTGAATAGTTCGGGCTTGTACACACCGCGCTGACACCA 1379  
 DB 1339 CGGATCAGCAATGTCCGGGTGAATAGTTCGGGCTTGTACACACCGCGCTGACACCA 1398

QY 1380 TGGAGTGGGTTTCCAGGAAGCAGTGTCTAACCGTAAAGAGGCGCTGCAACGCTG 1439  
 DB 1399 TGGAGTGGGTTTCCAGGAAGCAGTGTCTAACCGTAAAGAGGCGCTTACACGCTG 1458  
 QY 1440 AGATTGATGCTGGG 1455  
 DB 1459 TGGTCAATGACTTGGG 1474  
 RESULT 18  
 CQ796925 1467 bp DNA linear PAT 19-APR-2004  
 LOCUS Sequence 18 from Patent WO2004026772.  
 DEFINITION  
 CQ796925  
 VERSION CQ796925.1 GI:46408551  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 unclassified sequences.  
 REFERENCE  
 1 Hovanec, T.A.  
 AUTHORS Ammonia-oxidizing bacteria and methods of using and detecting  
 TITLE thesame  
 JOURNAL Patent: WO 2004026772-A 18 01-APR-2004;  
 Aquaria Inc. (US)  
 FEATURES  
 source Location/Qualifiers  
 1..1467  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"  
 /note="N. Aestuarii-like AOB P4clone42 16S rDNA"  
 ORIGIN  
 Query Match 93.2%; Score 1357.2; DB 6; Length 1467;  
 Best Local Similarity 97.2%; Pred. No. 2.9e-14;  
 Matches 1413; Conservative 0; Mismatches 38; Indels 3; Gaps 3;  
 QY 1 ATTGAACGCTGCGCGGACGCTTTTACACATGCAAGTGAACGCGACGACGCGATCTTGCAT 60  
 DB 16 ATTGAACGCTGCGCGGACGCTTTTACACATGCAAGTGAACGCGACGACGCGGCTTGCAC 75  
 QY 61 CTGTGGCGAGTGGCGGACGCGGTGAGTAATGCAATCGGAACGTAATCCAGAAAGAGGGGAGTA 120  
 DB 76 CTGTGGCGAGTGGCGGACGCGGTGAGTAATGCAATCGGAACGTAATCCAGAAAGAGGGGAGTA 135  
 QY 121 ACCGATGAAAGATGTGCTAATACCGCATTAATCTTAAGAGGAAAGCAGGGGATCGAA 180  
 DB 136 ACCGATGAAAGATGTGCTAATACCGCATTAATCTTAAGAGGAAAGCAGGGGATCGAA 195  
 QY 181 GACCTTGCCTTTTGGAGCGCGCGAGTGTCTGATTAAGTGTGTTGGGTTAAAGCCTTAC 240  
 DB 196 GACCTTGTGCTTTTGGAGCGCGCGAGTGTCTGATTAAGTGTGTTGGGTTAAAGCCTTAC 255  
 QY 241 CAAGGCGACATCAATGTTGTGCTGAGAGAGACAGACGACACTGGGACTGAGACAGC 300  
 DB 256 CAAGGCGACATCAATGTTGTGCTGAGAGAGACAGACGACACTGGGACTGAGACAGC 315  
 QY 301 GCCCAGACTCTTACGGGAGCAGAGTGGGAAATTTTGCACATGGGCGCAACGCTGATC 360  
 DB 316 GCCCAGACTCTTACGGGAGCAGAGTGGGAAATTTTGCACATGGGCGCAACGCTGATC 375  
 QY 361 CAGCAATGCGCGGTAGTGAAGAAGGCTTGGGTTTGAAGCTCTTTCACTGAGAGA 420  
 DB 376 CAGCAATGCGCGGTAGTGAAGAAGGCTTGGGTTTGAAGCTCTTTCACTGAGAGA 434  
 QY 421 AAAAGTTACGGTAAATATGTGACCATGACGGTATCGACAGAGAAGCAGCGCTTAAC 480  
 DB 435 AAAAGTTGATTAATATCAACTTATGACGTTACCGACAGAGAAGCAGCGCTTAAC 494  
 QY 481 TACGTCCAGCAGCCCGGTATACGTAAGGTGCAAGCGTTAAATCGAATTTACTGGCGT 540  
 DB 495 TACGTCCAGCAGCCCGGTATACGTAAGGTGCAAGCGTTAAATCGAATTTACTGGCGT 554







Qy 661 TGAATGCGTAGATATGAGAACATGATGCGGACGCTCTGCGTTAACT 720  
Db 681 TGAATGCGTAGATATGAGAACATGATGCGGACGCTCTGCGTTAACT 740  
Qy 721 GAGCTCATGACGAAAGCGTGGGAGCAACAGATTGATCCCTGGAGTCCAGCG 780  
Db 741 GAGCTCATGACGAAAGCGTGGGAGCAACAGATTGATCCCTGGAGTCCAGCG 800  
Qy 781 CTAAACGATGTCACACTAGTTGTTGGGCTTATAGCTTGGTGTACGAACTAAAGCGTGA 840  
Db 801 CTAAACGATGTCACACTAGTTGTTGGGCTTATAGCTTGGTGTACGAACTAAAGCGTGA 860  
Qy 841 AGTTGACCGGCTGGGAGTAGTCCGTCGACAGATTAACTCAAGAAATTGACGCGGACCC 900  
Db 861 AGTTGACCGGCTGGGAGTAGTCCGTCGACAGATTAACTCAAGAAATTGACGCGGACCC 920  
Qy 901 GCACAAAGCGGTGATATGATGATTAATTGATGCAACGCGGAAACCTTAACTACCTCT 960  
Db 921 GCACAAAGCGGTGATATGATGATTAATTGATGCAACGCGGAAACCTTAACTACCTCT 980  
Qy 961 GACATGAGCGAATTTCTAGATGATTAAGTCTTCCGGAACGCTAACACAGGTG 1017  
Db 981 GACATGAGCGAATTTCTAGATGATTAAGTCTTCCGGAACGCTAACACAGGTG 1040  
Qy 1018 CTGCAATGCTGCTGCTGAGTCTGTCGATGATGTTGGTTAACTCCGACGACGCA 1077  
Db 1041 CTGCAATGCTGCTGCTGAGTCTGTCGATGATGTTGGTTAACTCCGACGACGCA 1100  
Qy 1078 ACCCTGTCTATTAATGCTCATTTGTTGGGCACTTTAATGAGACTGCGCGGTGACAA 1137  
Db 1101 ACCCTGTCTATTAATGCTCATTTGTTGGGCACTTTAATGAGACTGCGCGGTGACAA 1160  
Qy 1138 CCGGAGGAGGTGGGAGTAGAGTCAAGTCCATGAGCTTAAAGGAGGCTTCAACG 1197  
Db 1161 CCGGAGGAGGTGGGAGTAGAGTCAAGTCCATGAGCTTAAAGGAGGCTTCAACG 1220  
Qy 1198 TAAATACAAATGCGCGTACAGAGGGTTGCCAACCCGCGAGGGAGCTAATCTCAGAAAGC 1257  
Db 1221 TAAATACAAATGCGCGTACAGAGGGTTGCCAACCCGCGAGGGAGCTAATCTCAGAAAGC 1280  
Qy 1258 GCGTCTGATGTCGAGTGTGAGTCTGCACTTCCGTGAAGTCGGAATTCGTAGTAAT 1317  
Db 1281 GCGTCTGATGTCGAGTGTGAGTCTGCACTTCCGTGAAGTCGGAATTCGTAGTAAT 1340  
Qy 1318 CCGGATCAGCATGTCGCGGTAATAGTCCCGGCTTGTATACACACCGCGCTCACAC 1377  
Db 1341 CCGGATCAGCATGTCGCGGTAATAGTCCCGGCTTGTATACACACCGCGCTCACAC 1400  
Qy 1378 CATGGAAGTGGTTTCAACAGAACAGTATCTTAACCGTAAGAGGCGCTTCCACGG 1437  
Db 1401 CATGGAAGTGGTTTCAACAGAACAGTATCTTAACCGTAAGAGGCGCTTCCACGG 1460  
Qy 1438 TGAGATTCATGACTGGGGGTG 1457  
Db 1461 TGAGATTCATGACTGGGGGTG 1480

RESULT 20  
AY123797 1499 bp DNA linear BCT 23-SEP-2003  
LOCUS Nitrosomonas sp. Nm84 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY123797  
VERSION AY123797.1 GI:24474426  
KEYWORDS Nitrosomonas sp. Nm84  
SOURCE Nitrosomonas sp. Nm84  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1 (bases 1 to 1499)  
AUTHORS Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roesser, A. and Koops, H.-P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal

of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
PUBMED 13130037  
REFERENCE 2 (bases 1 to 1499)  
AUTHORS Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roesser, A. and Wagner, M.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1..1499  
/organism="Nitrosomonas sp. Nm84"  
/mol\_type="genomic DNA"  
/strain="Nm84"  
/db\_xref="taxon:200124"  
<1..>1499  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 92.4%; Score 1346.2; DB 1; Length 1499;  
Best Local Similarity 95.5%; Pred. No. 4.3e-14;  
Matches 1394; Conservative 2; Mismatches 61; Indels 3; Gaps 1;  
Qy 1 ATTGAACGCTGGCGCATGCTTTTACATGCAAGTGCAGACGCGACGATGCTTGAT 60  
Db 3 ATTGAACGCTGGCGCATGCTTTTACATGCAAGTGCAGACGCGACGATGCTTGAT 62  
Qy 61 CTGCTGGCAGTGGCGGACGCGTGAAGTATGCAATCGGAAAGTATTCAGAAAGCGGAGTA 120  
Db 63 TNGTGGCAGTGGCGGACGCGTGAAGTATGCAATCGGAAAGTATTCAGAAAGCGGAGTA 122  
Qy 121 ACCGATCGAAAGTGTCTATATCCGCTATATCTTAAGAGAGAAAGCAGGGATCGAAA 180  
Db 123 ACCGATCGAAAGTGTCTATATCCGCTATATCTTAAGAGAGAAAGTATGAGATCGAAA 182  
Qy 181 GACCTTGGCGCTTTTGAAGGCGCGCATGCTGATTAAGTATGATGGGAGTAAAGCGCTAC 240  
Db 183 GACCTTATGCTTTTGAAGGCGCGCATGCTGATTAAGTATGATGGGAGTAAAGCGCTAC 242  
Qy 241 CAAGCGACGATCAGTATGTTGTTCTGAGAGAGACGACGACACTGGGACTGAGACAG 300  
Db 243 CAAGCGACGATCAGTATGTTGTTCTGAGAGAGACGACGACACTGGGACTGAGACAG 302  
Qy 301 GCCGAGCTCTTACGCGGACGACAGTGGGAAATTTTGGACAAATGGCGGAAAGCTGATC 360  
Db 303 GCCGAGCTCTTACGCGGACGACAGTGGGAAATTTTGGACAAATGGCGGAAAGCTGATC 362  
Qy 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCCGCTTGAAGCTCTTTCAGTGCAGAGA 420  
Db 363 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCCGCTTGAAGCTCTTTCAGTGCAGAGA 422  
Qy 421 AAAGTTACGTTAAATATGTAACCATGACGATATCGACAGAAAGACACCGGCTAAC 480  
Db 423 AAAGTTACGTTAAATATGTAACCATGTAATGTAATCGACAGAAAGACACCGGCTAAC 482  
Qy 481 TACGTGCACAGCCCGGTAATATCGATGAGTGAAGTGAATCCCGGGCTTAACTCGGGAGT 540  
Db 483 TACGTGCACAGCCCGGTAATATCGATGAGTGAAGTGAATCCCGGGCTTAACTCGGGAGT 542  
Qy 541 AAAGGTGCGCAGCGCGCTTGTAAAGTCAATGTAATCCCGGGCTTAACTCGGGAGT 600  
Db 543 AAAGGTGCGCAGCGCGCTTGTAAAGTCAATGTAATCCCGGGCTTAACTCGGGAGT 602  
Qy 601 TGCCTTGAATCAAAAGCTAGATGATGAGTGCAGAGGAGGTGAATTCATGTGTAGCAG 660  
Db 603 TGCCTTGAATCAAAAGCTAGATGATGAGTGCAGAGGAGGTGAATTCATGTGTAGCAG 662  
Qy 661 TGAATCGTAGAGATATGAGAAACATCGATGCGGACGACGCTCTGCTTAACTACT 720  
Db 663 TGAATCGTAGAGATATGAGAAACATCGATGCGGACGACGCTCTGCTTAACTACT 722  
Qy 721 GAGCTATACGAAAGCGTGGGAGCAACAGATTAATACCTGATGATGATGATGATGATGAT 780

Db 723 GACGCTCATGACGAAAGGGTGGGGAGCAACAGATAGATACCTGGTAGTCCAGCC 782  
Qy 781 CTAAACGATGTCACTAGTGTGGGCTTATTAGCTTGGTAGAGAACTAAACGGCTGA 840  
Db 783 CTAAACTAGTCACTAGTGTGGGCTTAAAGGCTTGGTAGTAAACGGCTGA 842  
Qy 841 AGTTGACCGCTGGGGAGTAGACGGTCCGAAGATTAAACCAAGAAATTGACGGGGACCC 900  
Db 843 AGTTGACCGCTGGGGAGTAGACGGTCCGAAGATTAAACCAAGAAATTGACGGGGACCC 902  
Qy 901 GCACAAAGCGGTGATTAATGATTAATTCGATGCAACGGGAAACCTTACCTACCTT 960  
Db 903 GCACAAAGCGGTGATTAATGATTAATTCGATGCAACGGGAAACCTTACCTACCTT 962  
Qy 961 GACATGTACGAAATTTTCTAGAGATAGATTAGT---CTTGGGAAACGCTAACACAGGTG 1017  
Db 963 GACATGTACGAAATTTTCTAGAGATAGATTAGT---CTTGGGAAACGCTAACACAGGTG 1022  
Qy 1018 CTGCAATGGCTGTGCTGAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAACGACGCA 1077  
Db 1023 CTGCAATGGCTGTGCTGAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAACGACGCA 1082  
Qy 1078 ACCCTGTGATTAAATTGCCATCATTTGTTGGGCACTTTAATGACATGCGGTGACAA 1137  
Db 1083 ACCCTGTGATTAAATTGCCATCATTTGTTGGGCACTTTAATGACATGCGGTGACAA 1142  
Qy 1138 CCGGAGGAAGTGGGGATGACGTCAGAGTCTCTATGGCCCTTATGGTTAGGGCTTCAACG 1197  
Db 1143 CCGGAGGAAGTGGGGATGACGTCAGAGTCTCTATGGCCCTTATGGTTAGGGCTTCAACG 1202  
Qy 1198 TAATCAATAGGCGCTGACAGAGGGTTCGCAACCCGAGGAGGAGTAACTCAGAAAC 1257  
Db 1203 TAATCAATAGGCGCTGACAGAGGGTTCGCAACCCGAGGAGGAGTAACTCAGAAAC 1262  
Qy 1258 GCGTGTAGTCCGATCGAGTCTGCACTCGATCGTGAATCGGAATCGTACTAAT 1317  
Db 1263 GCGTGTAGTCCGATCGAGTCTGCACTCGATCGTGAATCGGAATCGTACTAAT 1322  
Qy 1318 CCGGATCAGCATGTGCGCGGTGATACGTTCCGGGTCTTTGACACACCGCCCTGACAC 1377  
Db 1323 CCGGATCAGCATGTGCGCGGTGATACGTTCCGGGTCTTTGACACACCGCCCTGACAC 1382  
Qy 1378 CATGGAGTGGGTTTCAACGAGACGATAGTCTAACCGTAAAGAGGGGCTTGCACACG 1437  
Db 1383 CATGGAGTGGGTTTCAACGAGACGATAGTCTAACCGTAAAGAGGGGCTTGCACACG 1442  
Qy 1438 TGAGATTCTATGACTGGGGTG 1457  
Db 1443 TGAGATTCTATGACTGGGGTG 1462  
RESULT 21  
AY123798 1499 bp DNA linear BCT 23-SHP-2003  
LOCUS Nitrosomonas sp. Nm86 16S rDNA partial sequence.  
DEFINITION Nitrosomonas sp. Nm86 16S rDNA partial sequence.  
ACCESSION AY123798  
VERSION AY123798.1 GI:24474427  
KEYWORDS Nitrosomonas sp. Nm86  
SOURCE Nitrosomonas sp. Nm86  
ORGANISM Nitrosomonas sp. Nm86  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1 (bases 1 to 1499)  
AUTHORS Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roeser, A. and  
Koops, H.-P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
JOURNAL Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
PUBMED 1313037  
REFERENCE 2 (bases 1 to 1499)  
AUTHORS Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roeser, A. and  
Wagner, M.

TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1. 1499  
/organism="Nitrosomonas sp. Nm86"  
/mol\_type="genomic DNA"  
/strain="Nm86"  
/db\_xref="taxon:200125"  
<1. .>1499  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 92.3%; Score 1345.4; DB 1; Length 1499;  
Best Local Similarity 95.5%; Pred. NO. 4.4e-14;  
Matches 1395; Conservative 0; Mismatches 62; Indels 3; Gaps 1;  
Qy 1 ATTGAACGCTGGCGGATGCTTTACATGCAATGCAACGGCAGCAGATGCTTGCAT 60  
Db 3 ATTGAACGCTGGCGGATGCTTTACATGCAATGCAACGGCAGCAGATGCTTGCAT 62  
Qy 61 CTGTTGGGAGTGGCGGACGGGTGATGATGCAATGCAACGTTATCAGAAAGGGGGTA 120  
Db 63 CAGTGGCGAGTGGCGGACGGGTGATGATGCAATGCAACGTTATCAGAAAGGGGGTA 122  
Qy 121 ACGATCGAAAGATGCTTAATACCGCATATCTTAAGAGAGAAAGAGGGGATCGAAA 180  
Db 123 ACGATCGAAAGATGCTTAATACCGCATATCTTAAGAGAGAAAGAGGGGATCGAAA 182  
Qy 181 GACCTTGGCTTTTGGAGCGGCGCATGCTGATTAAGTGTGTTGGGTAAAGGCTTAC 240  
Db 183 GACCTTGGCTTTTGGAGCGGCGCATGCTGATTAAGTGTGTTGGGTAAAGGCTTAC 242  
Qy 241 CAAGCGACATCAATGTTGGTCTGAGAGAGACACAGCCACATGGGACTGAGACACG 300  
Db 243 CAAGCGACATCAATGTTGGTCTGAGAGAGACACAGCCACATGGGACTGAGACACG 302  
Qy 301 GCCGACCTCTTACGAGGAGGAGCGAGCTGGGGATTTTGGACAAATGGGCGCAAGCTGATC 360  
Db 303 GCCGACCTCTTACGAGGAGGAGCGAGCTGGGGATTTTGGACAAATGGGCGCAAGCTGATC 362  
Qy 361 CAGCATGCGCGGTGATGAGAAAGAGCCCTTGGGTGTTAAAGCTTTTCACTGAGAAAG 420  
Db 363 CAGCATGCGCGGTGATGAGAAAGAGCCCTTGGGTGTTAAAGCTTTTCACTGAGAAAG 422  
Qy 421 AAAGTTACGGTAATTAATCGTACCCATGACGGTATGACAGAAAGACACCGGCTAAC 480  
Db 423 AAAGTTACGGTAATTAATCGTACCCATGACGGTATGACAGAAAGACACCGGCTAAC 482  
Qy 481 TACGTGCCAGAGCGCGGATTAATAGTAGGGTGCAAGGCTTAATCGAATTAAGTGGGCT 540  
Db 483 TACGTGCCAGAGCGCGGATTAATAGTAGGGTGCAAGGCTTAATCGAATTAAGTGGGCT 542  
Qy 541 AAAGGTCGCGAGGCGGCTTGTAACTGATGATGTAATCCCGGGCTTAACTTGGGAAT 600  
Db 543 AAAGGTCGCGAGGCGGCTTGTAACTGATGATGTAATCCCGGGCTTAACTTGGGAAT 602  
Qy 601 TCGCTTTGAAACTACAAAGCTTAGAGTGGGAGAGAGGTGGAATTCATGTGTAGCAG 660  
Db 603 TCGCTTTGAAACTACAAAGCTTAGAGTGGGAGAGAGGTGGAATTCATGTGTAGCAG 662  
Qy 661 TGAATGGGTGAGATTAGAGAAACATGATGGCGAAGGAGGCTCTGAGGTTAACT 720  
Db 663 TGAATGGGTGAGATTAGAGAAACATGATGGCGAAGGAGGCTCTGAGGTTAACT 722  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATCCCTGATGTCACGCC 780  
Db 723 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATCCCTGATGTCACGCC 782  
Qy 781 CTAAACATGTCACTAGTGTGGGCTTTATGAGCTTGGTAACGAAGCTAACGGCTGCA 840  
Db 783 CTAAACATGTCACTAGTGTGGGCTTTATGAGCTTGGTAACGAAGCTAACGGCTGCA 842

841 AGTTGACCGCTGGGGAGTACGGTCGCAAGATTAAACTCAAGGAATTGACGGGAGCC 900  
843 AGTTGACCGCTGGGGAGTACGGTCGCAAGATTAAACTCAAGGAATTGACGGGAGCC 902  
901 GCAACAAGCGGTGATTTATGATTTATTCGATGCAAGCGGAAAACTTACCTACCTT 960  
903 GCACAAGCGGTGATTTATGATTTATTCGATGCAAGCGGAAAACTTACCTACCTT 962  
961 GACATGTACGAAATTTCTAGAGATAGATTAGTCT---TCGGGAACGCTTACACAGGTG 1017  
963 GACATGTACGAAATTTCTAGAGATAGATTAGTCTGCTCGAAAGAAATCTGAAACAGGTG 1022  
1018 CTGCAATGCTGTCTGCTACAGCTGTGTGTGATGATGTTGGTTAAAGTCCGCAACGACGCA 1077  
1023 CTGCAATGCTGTCTGCTACAGCTGTGTGTGATGATGTTGGTTAAAGTCCGCAACGACGCA 1082  
1078 ACCCTTGTCAATTAATTCGATCACTATTTGTTGGGCACTTTATATGACCTGCGGTGCAAA 1137  
1083 ACCCTTGTCAATTAATTCGATCACTATTTGTTGGGCACTTTATATGACCTGCGGTGCAAA 1142  
1138 CCGGAGGAAGTGGGGAGTACAGCTCAAGTCTCAATGACCCTTATGGGAGGCTTCAACAG 1197  
1143 CCGGAGGAAGTGGGGAGTACAGCTCAAGTCTCAATGACCCTTATGGGAGGCTTCAACAG 1202  
1198 TAATACATGCGCGCTGACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAAGC 1257  
1203 TAATACATGCGCGCTGACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAAGC 1262  
1258 GCGTGTAGTCCGAGATCGAGAGTCTGCACTGCACTCCGTGAAATCGGAAATCGTATGAAAT 1317  
1263 GCGTGTAGTCCGAGATCGAGAGTCTGCACTGCACTCCGTGAAATCGGAAATCGTATGAAAT 1322  
1318 CCGGATTCAGCAATGTCGGGTGAAATCGTCCGGGTCTGTATACACACCGCCGTCACAC 1377  
1323 CCGGATTCAGCAATGTCGGGTGAAATCGTCCGGGTCTGTATACACACCGCCGTCACAC 1382  
1378 CATGGAGTGGGTTTTCACCAAGAGCAGTAGTCTAACCCGTAAAGAGGCGCTTGCCACGG 1437  
1383 CATGGAGTGGGTTTTCACCAAGAGCAGTAGTCTAACCCGTAAAGAGGCGCTTGCCACGG 1442  
1438 TGAGATTCACTGCTGGGGTG 1457  
1443 TGAGATTCACTGCTGGGGTG 1462

RESULT 22  
ABI86828 1499 bp DNA linear ENV 11-AUG-2005  
LOCUS  
DEFINITION Uncultured bacterium gene for 16S rRNA, partial sequence,  
clone: TSAQ22.  
ACCESSION ABI86828  
VERSION ABI86828.1 GI:51100881  
KEYWORDS ENV.  
SOURCE uncultured bacterium  
ORGANISM uncultured bacterium  
Bacteria; environmental samples.

REFERENCE  
AUTHORS Yoshida, N., Takahashi, N. and Hiraishi, A.  
TITLE Phylogenetic characterization of a polychlorinated-dioxin-  
dechlorinating microbial community by use of microcosm studies  
JOURNAL Appl. Environ. Microbiol. 71 (8), 4325-4334 (2005)  
PUBMED 16085820  
REFERENCE 2 (bases 1 to 1499)  
AUTHORS Hiraishi, A. and Yoshida, N.  
TITLE Direct Submision  
JOURNAL Submitted (06-AUG-2004) Akira Hiraishi, Toyohashi University of  
Technology, Department of Ecological Engineering, Tempaku-cho,  
Toyohashi, Aichi, 441-8580, Japan (E-mail: hiraishi@eco.tut.ac.jp,  
Tel: 81-532-44-6913, Fax: 81-532-44-6929)  
FEATURES  
source  
1. .1499  
/organism="uncultured bacterium"  
/mol\_type="genomic DNA"

ORIGIN  
rRNA  
/isolation\_source="uncultured clone from polychlorinated  
dioxin dechlorinating microcosm"  
/db\_xref="taxon:77133"  
/clone="TSAQ22"  
/environmental\_sample  
<1..>1499  
/product="16S ribosomal RNA"

Query Match 91.9%; Score 1338.4; DB 3; Length 1499;  
Best Local Similarity 95.3%; Pred. No. 5.6e-14;  
Matches 1391; Conservative 0; Mismatches 66; Indels 3; Gaps 1;

1 ATTGAACGCTGGCGGCAATGCTTTATACATTCGACAGTGCAGAACCGGACACAGGATCTTGCAAT 60  
21 ATTGAACGCTGGCGGCAATGCTTTATACATTCGACAGTGCAGAACCGGACACAGGATCTTGCAAT 80  
61 CTGGTGGGAGTGGGGAGCGGGTGAATGATCGAAACGTATCCAGAAAGAGGGGGTGA 120  
81 TTGGTGGGAGTGGGGAGCGGGTGAATGATCGAAACGTATCCAGAAAGAGGGGGTGA 140  
121 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGCAGGGGATCGAAA 180  
141 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGCAGGGGATCGAAA 200  
181 GACCTTGGCGCTTTTGAAGCGCGCGATGTCTGATTAAGCTTGTGTTGGGTTAAAGGCTTAC 240  
201 GACCTTATCTTTTGAAGCGCGCGATGTCTGATTAAGCTTGTGTTGGGTTAAAGGCTTAC 260  
241 CAAGGCGAAGATCAGATGTTGTTGTTGAGAGAGCAGACCACTGAGACTGAGACAG 300  
261 CAAGGCGAAGATCAGATGTTGTTGTTGAGAGAGCAGACCACTGAGACTGAGACAG 320  
301 GCCAGACTCTTAACGGGAGCAGCAGTGGGGAATTTTGGCAATGGGCGCAAGCTTGATC 360  
321 GCCAGACTCTTAACGGGAGCAGCAGTGGGGAATTTTGGCAATGGGCGCAAGCTTGATC 380  
361 CAGCAATGCGCGGTGATGATGAAGAGCCCTTGGGTGTAAGCTCTTTCAGTTGAAGAAG 420  
381 CAGCAATGCGCGGTGATGATGAAGAGCCCTTGGGTGTAAGCTCTTTCAGTTGAAGAAG 440  
421 AAAGGTTAGCGTAAATTAATCGTGAACCATGACGGTATCGACAGAAAGACACCGGCTAAC 480  
441 AAAGGTTAGCGTAAATTAATCGTGAACCATGACGGTATCGACAGAAAGACACCGGCTAAC 500  
481 TACGTCCAGACCGCGGTAAATTAATCGTGAACCATGACGGTATCGACAGAAAGACACCGGCTAAC 540  
501 TACGTCCAGACCGCGGTAAATTAATCGTGAACCATGACGGTATCGACAGAAAGACACCGGCTAAC 560  
541 AAAGGTTAGCGTAAATTAATCGTGAACCATGACGGTATCGACAGAAAGACACCGGCTAAC 600  
561 AAAGGTTAGCGTAAATTAATCGTGAACCATGACGGTATCGACAGAAAGACACCGGCTAAC 620  
601 TGGCTTTGAAACTCAAGAGCTAGAGTGTGAGAGGAGGTGGAATTCATGTGTAGCAG 660  
621 TGGCTTTGAAACTCAAGAGCTAGAGTGTGAGAGGAGGTGGAATTCATGTGTAGCAG 680  
661 TGAATGCGTGAAGATATGAAAGATCGATGCGAAGGCGACCTCTCGGTTTAACTT 720  
681 TGAATGCGTGAAGATATGAAAGATCGATGCGAAGGCGACCTCTCGGTTTAACTT 740  
721 GACGCTCATGACGAAAGCGTGGGGAGCAAAACAGATTAATATCCCTGTATGTCACAGCC 800  
741 GACGCTCATGACGAAAGCGTGGGGAGCAAAACAGATTAATATCCCTGTATGTCACAGCC 820  
781 CTAAACGATTCGAACTGATGTTGTTGGGCTTATTAAGCTTGGTGAAGAGCTTACGCGTGA 840  
801 CTAAACGATTCGAACTGATGTTGTTGGGCTTATTAAGCTTGGTGAAGAGCTTACGCGTGA 860  
841 AGTTGACCGGCTGGGGAGTACGGTCCGACAGATTAAACTCAAGGAATTGACGGGAGCC 900  
861 AGTTGACCGGCTGGGGAGTACGGTCCGACAGATTAAACTCAAGGAATTGACGGGAGCC 920

QY 901 GCACAGCGGTGATTATGATGATTAATTCATGACGCAACGCAAACTTACTTACCTT 960  
| | | | |  
Db 921 GCACAGCGGTGATTATGATGATTAATTCATGACGCAACGCAAACTTACTTACCTT 980  
| | | | |  
QY 961 GACATGTACGAATTTCTAGAGATGATTAAGTCT--TCGGAAACGTTAACAGAGTG 1017  
| | | | |  
Db 981 GACATGTACGAATTTCTAGAGATGATTAAGTCTCGTCCGAAAGAACTTGAAACACAGTG 1040  
| | | | |  
QY 1018 CTGCATGCTGTCTGTACGCTCTGTGTGTGATGATGTGGTTAAGTCCCGCAACGACCGA 1077  
| | | | |  
Db 1041 CTGCATGCTGTCTGTACGCTCTGTGTGTGATGATGTGGTTAAGTCCCGCAACGACCGA 1100  
| | | | |  
QY 1078 ACCCTGTCAATTAATGACATCACTTGTGGCACTTTAATGACATGCGCGTGAACA 1137  
| | | | |  
Db 1101 ACCCTGTCAATTAATGACATCACTTGTGGCACTTTAATGACATGCGCGTGAACA 1160  
| | | | |  
QY 1138 CCGAGAGAGTGGGATGACGTCATGCTCAATGACCTTATGGGTAGGCTTCAACAG 1197  
| | | | |  
Db 1161 CCGAGAGAGTGGGATGACGTCATGCTCAATGACCTTATGGGTAGGCTTCAACAG 1220  
| | | | |  
QY 1198 TAATACATGCGCGCTGACAGAGGTTGCCAACCCCGAGGGGAGCTTAATTCAGAAAC 1257  
| | | | |  
Db 1221 TAATACATGCGCGCTGACAGAGGTTGCCAACCCCGAGGGGAGCTTAATTCAGAAAC 1280  
| | | | |  
QY 1258 GCGTGTAGTCCGAGATCGAGTCTGCACTCGACTCGTGAAGTGGGAATCGTATGAT 1317  
| | | | |  
Db 1281 GCGTGTAGTCCGAGATCGAGTCTGCACTCGACTCGTGAAGTGGGAATCGTATGAT 1340  
| | | | |  
QY 1318 CCGGATCAGCATGTGCGCGGTGAATCGTCCCGGCTTGTGACACACCGCCGTGACAC 1377  
| | | | |  
Db 1341 CCGGATCAGCATGTGCGCGGTGAATCGTCCCGGCTTGTGACACACCGCCGTGACAC 1400  
| | | | |  
QY 1378 CATGGAGTGGGTTTCAACGAAAGCAGGTAGTCTAACCGTAAAGAGGGGCTTTGCCACG 1437  
| | | | |  
Db 1401 CATGGAGTGGGTTTCAACGAAAGCAGGTAGTCTAACCGTAAAGAGGGGCTTTGCCACG 1460  
| | | | |  
QY 1438 TGAGATTCACTGACTGGGCTG 1457  
| | | | |  
Db 1461 TGAGATTCACTGACTGGGCTG 1480  
| | | | |  
RESULT 23  
AY343318 1534 bp DNA linear ENV 01-AUG-2005  
LOCUS AY343318  
DEFINITION Uncultured bacterium clone OSI 16S ribosomal RNA gene, partial  
sequence.  
ACCESSION AY343318  
VERSION AY343318.1 GI:37813565  
KEYWORDS ENV.  
SOURCE uncultured bacterium  
ORGANISM Bacteria; environmental samples.  
REFERENCE 1 (bases 1 to 1534)  
Sliemers,A.O., Haeyer,S., Schmid,M., Harhangi,H., Verwegen,K.,  
Kuenen,J.G. and Jetten,M.S.M.  
Nitrification and Anammox with Urea as the Energy Source  
TITLE Unpublished  
2 (bases 1 to 1534)  
Sliemers,A.O., Haeyer,S., Schmid,M., Harhangi,H., Verwegen,K.,  
Kuenen,J.G. and Jetten,M.S.M.  
Direct Submission  
JOURNAL Submitted (16-JUN-2003) Biotechnology, Delft University of  
Technology, Julianalaan 67, Delft 2628BC, The Netherlands  
FEATURES  
source  
1. 1534  
/organism="uncultured bacterium"  
/mol\_type="genomic DNA"  
/isolation\_source="batch reactor"  
/db\_xref="taxon:77133"  
/clone="OS1"  
/environmental\_sample  
<1..>1534  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 91.9%; Score 1338.4; DB 3; Length 1534;  
Best Local Similarity 95.3%; Pred. No. 5.5e-14;  
Matches 1391; Conservative 0; Mismatches 66; Indels 3; Gaps 1;  
QY 1 ATTTGACCTGCGGCATGCTTTTACATGCAAGTCGACCGGACGAGTGCCTTGAT 60  
| | | | |  
Db 21 ATTTGACCTGCGGCATGCTTTTACATGCAAGTCGACCGGACGAGTGCCTTGAT 80  
| | | | |  
QY 61 CTGTGGGAGTGGCGGACGGGTGATATGATCGTAACTGATCAAGAGGGGGTGA 120  
| | | | |  
Db 81 CTGTGGGAGTGGCGGACGGGTGATATGATCGTAACTGATCAAGAGGGGGTGA 140  
| | | | |  
QY 121 AGCATGCAAAAGTGTCTAATACCGCATATATCTTAAGAGGAAAGCAGGGGATCGAA 180  
| | | | |  
Db 141 AGCATGCAAAAGTGTCTAATACCGCATATATCTTAAGAGGAAAGTGGGATCGAA 200  
| | | | |  
QY 181 GACCTTGGCTTTTGGACGGCCGATGCTGATTAAGTGTGCTGAGGATTAAGGCTTAC 240  
| | | | |  
Db 201 GACCTTGGCTTTTGGACGGCCGATGCTGATTAAGTGTGCTGAGGATTAAGGCTTAC 260  
| | | | |  
QY 241 CAAGCGACGATCAGTATGATGCTGAGAGACGACACGCACTGGGACTGAGACAG 300  
| | | | |  
Db 261 CAAGCGACGATCAGTATGATGCTGAGAGACGACACGCACTGGGACTGAGACAG 320  
| | | | |  
QY 301 GCCCGACTCTTACCGGAGGACGAGTGGGAAATTTTGGACAAATGGGGGCAAGCCTGATC 360  
| | | | |  
Db 321 GCCCGACTCTTACCGGAGGACGAGTGGGAAATTTTGGACAAATGGGGGCAAGCCTGATC 380  
| | | | |  
QY 361 CAGCAATGCGCGTGTGAGAGAAAGCCTTGGGTTGAAAGCTTTTCACTGAGAGAA 420  
| | | | |  
Db 381 CAGCAATGCGCGTGTGAGAGAAAGCCTTGGGTTGAAAGCTTTTCACTGAGAGAA 440  
| | | | |  
QY 421 AAAGGTTACGGTAAATTAATCGTACCATGACGGTATCGACAGAAAGACACCGCTAAC 480  
| | | | |  
Db 441 AAAGGTTACGGTAAATTAATCGTACCATGACGGTATCGACAGAAAGACACCGCTAAC 500  
| | | | |  
QY 481 TAACGTCAGAGACCGCGGTAAATAGTGGGTGCAAGGCTTAATCGGAATTAAGTGGGCT 540  
| | | | |  
Db 501 TAACGTCAGAGACCGCGGTAAATAGTGGGTGCAAGGCTTAATCGGAATTAAGTGGGCT 560  
| | | | |  
QY 541 AAAGGTCGCGAGCGCGCTTGTAAAGTCAAGTGTGAAATCCCGGCTTAACTGGGAAT 600  
| | | | |  
Db 561 AAAGGTCGCGAGCGCGCTTGTAAAGTCAAGTGTGAAATCCCGGCTTAACTGGGAAT 620  
| | | | |  
QY 601 TCGCTTGAATCTAAGAGTGTGCAAGAGGAGTGTGAATTCATGTGTAGCAG 660  
| | | | |  
Db 621 TCGCTTGAATCTAAGAGTGTGCAAGAGGAGTGTGAATTCATGTGTAGCAG 680  
| | | | |  
QY 661 TGAATATGCTAGATATGGAAGAACATCGATGGCGGAAGGCTCTGGGTTAAAGCT 720  
| | | | |  
Db 681 TGAATATGCTAGATATGGAAGAACATCGATGGCGGAAGGCTCTGGGTTAAAGCT 740  
| | | | |  
QY 721 GACGCTCATGACGAAACGCTGGGAGCAACAGAGTTAGATACCTGTTAGTCCAGCC 780  
| | | | |  
Db 741 GACGCTCATGACGAAACGCTGGGAGCAACAGAGTTAGATACCTGTTAGTCCAGCC 800  
| | | | |  
QY 781 CTAAACGATGTCACTAGTGTGGCTTATTAAGCTTGTGAACGAAGCTAACCGCTGA 840  
| | | | |  
Db 801 CTAAACGATGTCACTAGTGTGGCTTATTAAGCTTGTGAACGAAGCTAACCGCTGA 860  
| | | | |  
QY 841 AGTTGACCGCTGGGAGATAGGTGTGCAAGATTAATACTCAAGAGAAATGACGGGAGACC 900  
| | | | |  
Db 861 AGTTGACCGCTGGGAGATAGGTGTGCAAGATTAATACTCAAGAGAAATGACGGGAGACC 920  
| | | | |  
QY 901 GCACAGCGGTGATTAATGATGATTAATTCATGACGCAACGCAAACTTACTTACCTT 960  
| | | | |  
Db 921 GCACAGCGGTGATTAATGATGATTAATTCATGACGCAACGCAAACTTACTTACCTT 980  
| | | | |  
QY 961 GACATGTACGAATTTCTAGAGATGATTAAGTCT--TCGGAAACGTTAACAGAGTG 1017  
| | | | |  
Db 981 GACATGTACGAATTTCTAGAGATGATTAAGTCTCGTCCGAAAGAACTTGAAACACAGTG 1040  
| | | | |

QY	1018	UTGCATGGCTGCTCGTCAAGTCGTGTCGAGATAGTGTGGGTTAACTCCCGAAGAGGCGCA	1077
Db	1041	CTGCATGGCTGTGTCAGCTCGTGTCCGAGATGTGGGTTAAAGTCCCGAAGAGCGCA	1100
QY	1078	ACCCCTTGTCATTAATTCGCATCATTTGGTTGGGCACTTTAATGAGACTGCCGATGACAAA	1137
Db	1101	ACCCCTTGTCATTAATTCGCATCATTTGGTTGGGCACTTTAATGAGACTGCCGATGATAAA	1160
QY	1138	CCGAGAGAAAGTGGGGATGACGTCAAGTCTTCATAGGCCCTTAATGGGTAAAGGCTTCACAG	1197
Db	1161	CCGAGAGAAAGTGGGGATGAGACTCAAGTCTTCATAGGCCCTTAATGGGTAAAGGCTTCACAG	1220
QY	1198	TAAATCAATGGCGCGGTACAGAGGGTTGGTCCAAACCGCGAGGGGAGACTAATTCAGAAAAGC	1257
Db	1221	TAAATCAATGGCGCGGTACAGAGGGTTGGTCCAAACCGCGAGGGGAGACTAATTCAGAAAAGC	1280
QY	1258	GCGTCGTAGTCCGATTCGGAAGTCTTGCAACTCGACTCGGTGAAGTCGGAAATCGTAAGTAAT	1317
Db	1281	GCGTCGTAGTCCGGAATCGGAATCTTGCAACTCGACTCGGTGAAGTGGAAATGCTAAGTAAT	1340
QY	1318	CGCGGATAGCATGTGCGCGGTGAATACGTTCCCGGGCTTGTAAACAACCGCCGCTCACAC	1377
Db	1341	CGCGGATAGCATGTGCGCGGTGAATACGTTCCCGGGCTTGTAAACAACCGCCGCTCACAC	1400
QY	1378	CATGGAGATGGGTTTCAACAGAAAGCAGGTAGTCTAAACCGTAAGAGAGGCGCTTGCCACGG	1437
Db	1401	CATGGAGATGGGTTTCAACAGAAAGCAGGTAGTCTAAACCGCAAGAGGCGCTTGCCACGG	1460
QY	1438	TGAGATTCAATGACTGGGGTG 1457	
Db	1461	TGAGGTTCAATGACTGGGGTG 1480	

RESULT 24			
AB117705	1486 bp	DNA	linear
LOCUS			ENV 27-JUL-2005
DEFINITION	Uncultured bacterium gene for 16S ribosomal RNA, partial sequence, clone:NB-01.		

SOURCE	ORGANISM	REFERENCE
uncultured bacterium	uncultured bacterium	1
Bacteria; environmental samples.	Bacteria; environmental samples.	Kindaichi, T., Ito, T. and Okabe, S.
Ecophysiological interaction between nitrifying bacteria and	Ecophysiological interaction between nitrifying bacteria and	

JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Appl. Environ. Microbiol. 70 (3), 1641-1650 (2004)  
1506789  
2 (bases 1 to 1486)  
Kindaichi, T.  
Direct Submission  
Submitted (18-AUG-2003) Tomonori Kindaichi, Hokkaido university.  
graduate school of engineering, department of urban and  
environmental engineering, Kita 13, nishi 8, Kita-ku, Sapporo,  
Hokkaido, 060-8628, Japan (E-mail: kindaichi.tokuaki@ipc.hokudai.ac.jp,  
Tel:81-11-706-6267, Fax:81-11-706-6267)

```

FEATURES
source      location/Qualifiers
1. 1486
/organism="uncultured bacterium"
/mol_type="genomic DNA"
/isolation_source="autotrophic nitrifying biofilm"
/db_xref="taxon:77133"
/clone="NB-01"
/environmental_sample
<1. >1486
/product="16S ribosomal RNA"

```

Query Match	91.6%	Score 1335.2;	DB 3;	Length 1486;
Best Local Similarity	97.9%;	Pred. No. 6.4e-14;		
Matches 1426;	Conservative	0;	Mismatches 24;	Indels 7;
			Gaps	7;

QY	1	ATTGAAAGCTGGGGGAGATGCTTTACACATGCAAGTGAACCGGACGACGGATCTTGAT	60
Db	18	ATTGAAAGCTGGGGGAGATGCTTTACACATGCAAGTGAACCGGACGACGGATGCTTGAT	77
QY	61	CTGGTGGCGATGGCGGACGGGTGAAGTATGCAATCGAAAGTATCCAAAGAGCGGGGTA	120
Db	78	TTGGTGGCGATGGCGGACGGGTGAAGTATCAATCGAAAGTATCCAAAGAGCGGGGTA	137
QY	121	ACGCATCGAAAGATGTGCTTAATCCGCATTAATCTTAAGAGGAAGACGGGGATGAAA	180
Db	138	ACGCATCGAAAGATGTGCTTAATCCGCATTAATCTTAAGAGGAAGACGGGGATGAAA	197
QY	181	GACCTTGCGCTTTTGGAGCGCGGATGTCTGATTTAGCTAGTTGGTGGGGTAAAGGCTAC	240
Db	198	GACCTTGCGCTTTTGGAGCGCGGATGTCTGATTTAGCTAGTTGGTGGGGTAAAGGCTAC	257
QY	241	CAAGCGCAGCATGATGATTGGTCTGAGAGACGACACGCTACACTGTGGACTGAGACG	300
Db	258	CAAGCGCAGCATGATGATTGGTCTGAGAGACGACACGCTACACTGTGGACTGAGACG	317
QY	301	GCCCAAGCTCTCTACGGGAGCAGACATGGGGAAATTTTGGACAAATGGCCGCAAGCTGATC	360
Db	318	GCCCAAGCTCTCTACGGGAGCAGACATGGGGAAATTTTGGACAAATGGCCGCAAGCTGATC	377
QY	361	CAGCAATGCCGCTGATGATGAAGAAAGGCTTCGGGTGTGAAGCTCTTTCAGTCAGAGA	420
Db	378	CAGCAATGCCGCTGATGATGAAGAAAGGCTTCGGGTGTGAAGCTCTTTCAGTCAGAGA	437
QY	421	AAAGCTTACGGTAAATATATCTGATCCATACCGTATCCACAGAAAGAGACGGGGCTAAC	480
Db	438	AAAGGCTGCTGTAAATTAATCT-GCAGCTCATATACGGTATCCACAGAAAGAGACGGGGCT-AAC	495
QY	481	TAGCTGCAGAGCGCGGTAATATCGTAGGGTGCAGACGGTAAATCGGAATTAATCTGGGCGT	540
Db	496	TAGCTGCAGAGC-CGCGGTAATACGTA-GGTGCAGCGTAAATCGGAATTAATCTGGGCGT	553
QY	541	AAAGGATGCCGAGCGGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAT	600
Db	554	-AAAGGATGCCGAGCGGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAT	612
QY	601	TGCGTTTGAACCTTCAAAAGCTAGTGTGGCAGAGGGAGGTGGAAATCCAAATGTATACAG	660
Db	613	TGCGTTTGAACCTTCAAAAGCTAGTGTGGCAGAGGGAGGTGGAAATCCAAATGTATACAG	672
QY	661	TGAATCGATGAGATATGAGAAACATCGATGCGAGGACAGGCTCTGGGTTAACACT	720
Db	673	TGAATCGATGAGATATGAGAAACATCGATGCGAGGACAGGCTCTGGGTTAACACT	732
QY	721	GAGCTCATGACAGAAAGCGTGGGAGCAACAGGATTAAGTATCCCTGTGTATCCAGGCC	780
Db	733	GAGCTCATGACAGAAAGCGTGGGAGCAACAGGATTAAGTATCCCTGTGTATCCAGGCC	792
QY	781	CTAAACGATGTCACACTAGTTGTGGGCTTAATTAAGGCTTGGTATACGAAGCTAACGCGTGA	840
Db	793	CTAAACGATGTCACACTAGTTGTGGGCTTAATTAAGGCTTGGTATACGAAGCTAACGCGTGA	852
QY	841	AGTTGACCGCTGGGGAGTACGGTTCGCAAGTTAAACTCAAGAGAAATTCGCGGAGACC	900
Db	853	AGTTGACCGCTGGGGAGTACGGTTCGCAAGTTAAACTCAAGAGAAATTCGCGGAGACC	912
QY	901	GCAACAAGCGGTGATATATGTGGATTAATTCGATCCAAACGCGAAGAAACCTTACCTACCTT	960
Db	913	GCAACAAGCGGTGATATATGTGGATTAATTCGATCCAAACGCGAAGAAACCTTACCTT	971
QY	961	GACATGTAGCAATTTTCTAGAGATAGATTAGTCTTCGGGAGCGCTTAACACAGGTGCTG	1020
Db	972	GACATGTAGCAATTTTCTAGAGATAGATTAGTCTTC-CGGAAGCTTAACACAGGTGCTG	1030
QY	1021	CATGCTGTCTCAAGCTCTGTCTCGTAGATTTGGGTTAAAGTCCCGCAGACGAGCGCAACC	1080

Db 1031 CATGGCTGCTAGCTCGTCTGTAAGTGTGGTTAAGTCCCGCAACGAGCGCAAC 1090  
Qy 1081 CTGTTCATTAAATTTGATCATATTTGGTTGGGCACTTTAATGAGACTGCCGATGCAAAACCG 1140  
Db 1091 CTGTTCATTAAATTTGATCATATTTAAGTGGGCACTTTAATGAGACTGCCGATGCAAAACCG 1150  
Qy 1141 GAGGAAGTGGGGAATGACGTCAAGTCTCATAGGCCCTTATGGTAGGGCTTCAACAGTAA 1200  
Db 1151 GAGGAAGTGGGGAATGACGTCAAGTCTCATAGGCCCTTATGGTAGGGCTTCAACAGTAA 1210  
Qy 1201 TACAATGGCGCGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAACGGCG 1260  
Db 1211 TACAATGGCGCGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAACGGCG 1270  
Qy 1261 TCGTAGTCCGAGTCCGAGTCTGCACTCGACTCGTGAAGTCGGAATCGCTAGTAATCCG 1320  
Db 1271 TCGTAGTCCGAGTCCGAGTCTGCACTCGACTCGTGAAGTCGGAATCGCTAGTAATCCG 1330  
Qy 1321 GGATCAGCAGTCTCGCGGTAAATACGTTCCCGGCTCTTGTACACAGCCCGCTCAGACCAT 1380  
Db 1331 GGATCAGCAGTCTCGCGGTAAATACGTTCCCGGCTCTTGTACACAGCCCGCTCAGACCAT 1390  
Qy 1381 GGGAGTGGGTTTCCACGAGAGCAGGTAGTCTAACCGTAAAGAGGGCGCTTGCCACGGTGA 1440  
Db 1391 GGGAGTGGGTTTCCACGAGAGCAGGTAGTCTAACCGTAAAGAGGGCGCTTGCCACGGTGA 1450  
Qy 1441 GATTCTAGACTGGGGTG 1457  
Db 1451 GATTCTAGACTGGGGTG 1467

RESULT 25  
LOCUS AJ621027 1454 bp DNA linear BCT 08-JUN-2004

DEFINITION Nitrosomonas sp. Is32 16S rRNA gene, isolate Is32.

ACCESSION AJ621027 1. GI:40994846

VERSION 16S ribosomal RNA, 16S rRNA gene.

KEYWORDS Nitrosomonas sp. Is32

SOURCE Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;

ORGANISM Nitrosomonadaceae; Nitrosomonas.

REFERENCE 1 Nielsen, S., Baer Gilleisen, M.J., Revsbech, N., Laanbroek, H.J. and

AUTHORS Bollmann, A.

TITLE Influence of heterotrophic bacteria on the activity and growth

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1454)

AUTHORS Bollmann, A.

TITLE Direct Submission

JOURNAL Submitted (15-JAN-2004) Bollmann A., Department for Microbial

ECOLGY, NIOO-KNAW Centre for Limnology, Rijksstraatweg 6, 3631 AC

NEUWERELUIS, NETHERLANDS

FEATURES

source 1.1454

location/Qualifiers

1.1454

/organism="Nitrosomonas sp. Is32"

/mol\_type="genomic DNA"

/isolate="Is32"

/isolation\_source="fresh water sediment"

/db\_xref="taxon:261293"

/country="Netherlands;Lake Drontemeer"

1.1454

/gene="16S rRNA"

1.1454

/gene="16S rRNA"

/product="16S ribosomal RNA"

ORIGIN

Query Match 91.6%; Score 1334.6; DB 1; Length 1454;

Best Local Similarity 95.3%; Pred. No. 6.7e-14;

Matches 1386; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

Qy 1 ATTGAACGTGGCGGCATGCTTTACATGCAAGTCSAAACGGCAGCAGGATGCTTGCAAT 60  
Db 1 ATTGAACGTGGCGGCATGCTTTACATGCAAGTCSAAACGGCAGCAGGATGCTTGCAAT 60  
Qy 61 CTGTGGGAGTGGCGGAGCGGGTGAATATGATCGAAGCTATTCAGAAAGGGGGGTA 120  
Db 61 CTGTGGGAGTGGCGGAGCGGGTGAATATGATCGAAGCTATTCAGAAAGGGGGGTA 120  
Qy 121 ACGCATCGAAGATGTCTAATACCGCATATCTCTTAAGAGGAAAGCAGGGGATTCGAAA 180  
Db 121 ACGCATCGAAGATGTCTAATACCGCATATCTCTTAAGAGGAAAGCAGGGGATTCGAAA 180  
Qy 181 GACCTTGGCTTTTGAAGGGCGGATGCTGATTAAGTGAATGTTGGGTAAAGGCTTAC 240  
Db 181 GACCTTGGCTTTTGAAGGGCGGATGCTGATTAAGTGAATGTTGGGTAAAGGCTTAC 240  
Qy 241 CAGGCGCAGATCAGTATGTTGTTGAGAGGACGACAGCCACACTGGGACTGAGACAG 300  
Db 241 CAGGCGCAGATCAGTATGTTGTTGAGAGGACGACAGCCACACTGGGACTGAGACAG 300  
Qy 301 GCCGAGACTCTTACCGGAGGCGAGCAGTGGGGAATTTGACAAATGGCGCAGAGCTGATC 360  
Db 301 GCCGAGACTCTTACCGGAGGCGAGCAGTGGGGAATTTGACAAATGGCGCAGAGCTGATC 360  
Qy 361 CAGCAATGCGCGGTGAGTGAAGAAAGGCGCTGGGTTGTAAGCTCTTCAAGTCAGAGA 420  
Db 361 CAGCAATGCGCGGTGAGTGAAGAAAGGCGCTGGGTTGTAAGCTCTTCAAGTCAGAGA 420  
Qy 421 AAAAGTTACGGTAAATTAATCGTACCATGACCGTATCGACAGAGAAAGCAACCGCTAAC 480  
Db 421 AAAAGTTACGGTAAATTAATCGTACCATGACCGTATCGACAGAGAAAGCAACCGCTAAC 480  
Qy 481 TACGTCAGCAGACGCCGGTAAATACGTAGGGTGCAGAGCTTAATCGGAATTAATCGGGGT 540  
Db 481 TACGTCAGCAGCAGCGCGGTAAATACGTAGGGTGCAGAGCTTAATCGGAATTAATCGGGGT 540  
Qy 541 AAAAGGTGGGAGGCGGCTTGTAGTGAATGTGAATTCCTGGGCTTAACTTGGGAAAT 600  
Db 541 AAAAGGTGGGAGGCGGCTTGTAGTGAATGTGAATTCCTGGGCTTAACTTGGGAAAT 600  
Qy 601 TCGCTTTGAAATCTACAAAGCTAGAGTGGGAGGAGGAGTGGAAATTCATGTGTAGCAG 660  
Db 601 TCGCTTTGAAATCTAGAGTGGGAGGAGGAGTGGAAATTCATGTGTAGCAG 660  
Qy 661 TGAATGCGTGAAGATTAAGGAAGAAATCGATGCGGAAAGCGAGCTCTCGGTTAACT 720  
Db 661 TGAATGCGTGAAGATTAAGGAAGAAATCGATGCGGAAAGCGAGCTCTCGGTTAACT 720  
Qy 721 GACGCTCAGTACGAGAAAGCGTGGGAGGAGCAAGATTAGATCCCTGTAGTCCAGCGC 780  
Db 721 GACGCTCAGTACGAGAAAGCGTGGGAGGAGCAAGATTAGATCCCTGTAGTCCAGCGC 780  
Qy 781 CTAAACGATGTCACATAGTGTGGGCTTATTAAGCTTGTGAACGAGTAAACGCGTGA 840  
Db 781 CTAAACGATGTCACATAGTGTGGGCTTATTAAGCTTGTGAACGAGTAAACGCGTGA 840  
Qy 841 AGTTAACCGGCTGGGAGTACGCTGCAAGATTAAATCTCAAGGAAATTTGACGGGAGACC 900  
Db 841 AGTTAACCGGCTGGGAGTACGCTGCAAGATTAAATCTCAAGGAAATTTGACGGGAGACC 900  
Qy 901 GCACAAAGGCGTGAATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTCT 960  
Db 901 GCACAAAGGCGTGAATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTCT 960  
Qy 961 GACATGTAGCAAAATTTTCAAGAGTGAATTTGTGCGGAAAGGAAATCGAACAGGTG 1020  
Db 961 GACATGTAGCAAAATTTTCAAGAGTGAATTTGTGCGGAAAGGAAATCGAACAGGTG 1020  
Qy 1018 CTGCAATGCTGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
Db 1021 CTGCAATGCTGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

QY	1078	ACCCCTGTGCATTAATTCGCATCATTTGGTTGGCACTTTTAATGAGACTGCCGGTGACAA	1137
Db	1081	ACCCCTGTGCATTAATTCGCATCATTTAGTTGGCACTTTATGAGACTGCCGGTGATMAA	1140
QY	1138	CCGGAAGAAAGGTGGGGGANGAGCTGAAGCTCATGGCCCTTAATGGTAGGGCTTCACACG	1197
Db	1141	CCGGAAGAAAGGTGGGGGANGAGCTGAAGCTCATGGCCCTTAATGGTAGGGCTTCACACG	1200
QY	1198	TAATACAATGCGCGGTACAGAGGGGTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGC	1257
Db	1201	TAATACAATGCGCGGTACAGAGGGGTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGC	1260
QY	1258	GCGTGTGATGTCGCGATCGGAGTCTTGCAACTTGACTCCGTGAAGTCGGAATTCGATGAT	1317
Db	1261	GCGTGTGATGTCGCGATCGGAGTCTTGCAACTTGACTCCGTGAAGTCGGAATTCGATGAT	1320
QY	1318	CGCGGATACAGATGTCGCGGTGAATATAGCTCCGGGTCTTGATACACACCGCCGCTACAC	1377
Db	1321	CGCGGATACAGATGTCGCGGTGAATATAGCTCCGGGTCTTGATACACACCGCCGCTACAC	1380
QY	1378	CATGGAGATGGGTTTCAACCAAGACGAGTAGTCTTAACCGTAAGAGGCGCGTTCACACG	1437
Db	1381	CATGGAGATGGGTTTCAACCAAGACGAGTAGTCTTAACCGTAAGAGGAGCGTTCACACG	1440
QY	1438	TGAGATTCATGACT 1451	
Db	1441	TGAGATTCATGACT 1454	
RESULT 26			
LOCUS	AY123811	1499 bp	linear BCT 23-SEP-2003
DEFINITION	Nitrosomonas sp. Nm59 16S ribosomal RNA gene, partial sequence.		
ACCESSION	AY123811		
VERSION	AY123811.1 GI:24474440		
KEYWORDS			
SOURCE	Nitrosomonas sp. Nm59		
ORGANISM	Nitrosomonas sp. Nm59		
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosomonas.		
AUTHORS	1 (bases 1 to 1499) Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Rozer,A. and Koops,H.P.		
TITLE	16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal of a new lineage within the nitrosomonads		
JOURNAL	Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)		
PUBMED	13130037		
REFERENCE	2 (bases 1 to 1499) Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Roesser,A. and Wagner,M.		
AUTHORS	Direct Submission Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany		
TITLE	Location/Qualifiers		
FEATURES	1..1499		
source	/organism="Nitrosomonas sp. Nm59" /mol_type="genomic DNA" /strain="Nm59" /db_xref="taxon:200127" <1..>1499 /product="16S ribosomal RNA"		
ORIGIN			
Query Match	91.5%;	Score 1332.6;	DB 1; Length 1499;
Best Local Similarity	95.0%;	Pred. No. 6.9e-14;	
Matches 1387;	Conservative	0; Mismatches	70; Indels 3; Gaps 1
QY	1	ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAACGGCAGACCGATGCTTGCA	60
Db	3	ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAACGGCAGACCGATGCTTGCA	62
QY	61	CTGGTGCCAGATGCGCGACGGGTGAGTAATGCATCGGAACGTAATCCAGAAAGGGGGGTA	120

Db	63	TTNGTGGCAAGTGGGGAAACGGGTGATTAATGCATCGGAACGTGCTTAAAGTGGGGGATTA	1221
Qy	121	ACGCATCGAAAGATGTGCTAATACCGCATATACCTCTAAGAGAAACACAGGGGATCGAAA	180
Db	123	ACGCATCGAAAGATGTGCTAATACCGCATATATCTCTCAGAGGAGAAATGAGGGATCGAAA	182
Qy	181	GACCTTGCCCTTTTGGAGCGGCCATGCTGATTAAGCTAGTGGTGGGGTAAAGGCTTAC	240
Db	183	GACCTTATCTTTAAGAGCGGCCATGTCTGATTAAGCTAGTGGTGGAGTAAATGGCTTAC	242
Qy	241	CAAGCGACGATCAGTAGTTGGTCTGAGAGGACGACGACACTGAGACG	300
Db	243	CAAGCGACGATCAGTAGTTGGTCTGAGAGGACGACGACACTGAGACG	302
Qy	301	GCCCAAGCTCTTACCGGAGGCGACAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTTATC	366
Db	303	GCCCAAGCTCTTACCGGAGGCGACAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTTATC	362
Qy	361	CACCAATGCCGCGTGAAGTGAAGAAGCCCTGGGGTGTAAAGCTCTTCAAGTCGAGAAAGA	420
Db	363	CACCAATGCCGCGTGAAGTGAAGAAGCCCTGGGGTGTAAAGCTCTTCAAGTGAAGAGA	422
Qy	421	AAAGGTTACGGTAAATTAATCGTGACCCATGACGGTATCGACAGAAAGACCGGCTTAC	480
Db	423	AAATAATGATCTAATATACAAATTTGATGACGGTATCAACAGAAAGACCGGCTTAC	482
Qy	481	TACGTGCGACAGCCGCGGTAAATACGTAGGGTGCAGCGTTAATCGGAATTAATCTGGGCGT	546
Db	483	TACGTGCGACAGCCGCGGTAAATACGTAGGGTGCAGCGTTAATCGGAATTAATCTGGGCGT	542
Qy	541	AAAGGGTGCAGAGCGGCGCTTGTATGATGACAGATGTAATCCCGGGCTTAACTGGGAAT	600
Db	543	AAAGGGTGCAGAGCGGCGCTTGTATGATGACAGATGTAATCCCGGGCTTAACTGGGAAT	602
Qy	601	TGCGTTTGAATCTAACAAGCTAGAGTGTGGCAGAGGGAGTGGAAATTCATGTGTAGCAG	666
Db	603	TGCGTTTGAATCTAACAAGCTAGAGTGTGGCAGAGGGAGTGGAAATTCATGTGTAGCAG	662
Qy	661	TGAATCCGTAGATATGAGAAAGACATGATGGCGAAAGCGACCTCCTGGGTTAAACAT	722
Db	663	TGAATCCGTAGATATGAGAAAGACATGATGGCGAAAGCGACCTCCTGGGTTAAACAT	722
Qy	721	GACGCTATCAAGAAAGGTGGGGAGCAAAACAGGATTAAGATACCCTGTAGTCCACGCC	780
Db	723	GACGCTATCAAGAAAGGTGGGGAGCAAAACAGGATTAAGATACCCTGTAGTCCACGCC	782
Qy	781	CTAAACGATGTCAACTAGTGTGTGGGCTTATTAAGCTTGTGTAAGCAAGCTAACGCGTGA	840
Db	783	CTAAACGATGTCAACTAGTGTGTGGGCTTATTAAGCTTGTGTAAGCTAACGCGTGA	842
Qy	841	AGTTGACCGGCTGGGGAGTACGGTGCAGAAATTTAAACTCAAGGAATTTGACGGGGACCC	900
Db	843	AGTTGACCGGCTGGGGAGTACGGTGCAGAAATTTAAACTCAAGGAATTTGACGGGGACCC	902
Qy	901	GCAACAACCGGTGATTAATGGAATTAATTCGATNGACAACGCGAAAACCTTACCTTACCTT	966
Db	903	GCAACAACCGGTGATTAATGGAATTAATTCGATNGACAACGCGAAAACCTTACCTTACCTT	962
Qy	961	GACATGTAGGGAATTTTCTAGAGATAGATTAGTG--CTTCCGGAAACGCTAACACAGGTG	1020
Db	963	GACATGTAGGGAATTTTCTAGAGATAGATTAGTG--CTTCCGGAAACGCTAACACAGGTG	1022
Qy	1018	CTGCATGGCTGTGCTCAGCTCGTGTGCTGAGATTTGGGTTAAGTCCGCAACGAGCCCA	1080
Db	1023	CTGCATGGCTGTGCTCAGCTCGTGTGCTGAGATTTGGGTTAAGTCCGCAACGAGCCCA	1082
Qy	1078	ACCTTGTCATTAATGGCCATCAATTTGGTGTGGGCACTTAAATGAGACGTCGGGTGACAA	1140
Db	1083	ACCTTGTCATTAATGGCCATCAATTTAGTTGGGCACTTAAATGAGACGTCGGGTGACAA	1142
Qy	1138	CCGAGAGAGGTGGGAGTGAAGTCAAGTCTCAATGGCCCTTAATGGGTAGGGCTTACAG	1200



Db 1143 CCGAGGAAGTGGGGATGACGTCAAGTCTCATGSCCTTAATGGGTAGGGCTTCACAG 1202  
Qy 1198 TAAATCAATGGCCGCTACAGAGGTTGCCAACCCGGAGGGAGGTAATCTCAGAAGC 1257  
Db 1203 TAAATCAATGGCCGCTACAGAGGTTGCCAACCCGGAGGGAGGTAATCTCAGAAGC 1262  
Qy 1258 GCGTGTAGTCCGGATCGAGTCTGCAACTCGACTCCGTGAAGTCGGAATCCGTAGTAAT 1317  
Db 1263 GCGTGTAGTCCGGATCGAGTCTGCAACTCGACTCCGTGAAGTCGGAATCCGTAGTAAT 1322  
Qy 1318 CCGGATCAGCATGTGCGGGTGAATACGTTCCGGGTCTTTGTAACACACCGCCGCTACAC 1377  
Db 1323 CCGGATCAGCATGTGCGGGTGAATACGTTCCGGGTCTTTGTAACACACCGCCGCTACAC 1382  
Qy 1378 CATGGAGTGGGTTTACCAAGACAGTATGTTACCGTAAGAGGGGCTTTGCCACCG 1437  
Db 1383 CATGGAGTGGGTTTACCAAGACAGTATGTTACCGTAAGAGGGGCTTTGCCACCG 1442  
Qy 1438 TGAGATTGATGACTGGGGTG 1457  
Db 1443 TGAGATTGATGACTGGGGTG 1462

RESULT 27  
AY123810 1499 bp DNA linear BCT 23-SBP-2003  
LOCUS Nitrosomonas sp. Nm47 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY123810  
ACCESSION AY123810.1 GI:24474439  
VERSION  
KEYWORDS  
SOURCE Nitrosomonas sp. Nm47  
ORGANISM Nitrosomonas sp. Nm47  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
REFERENCE  
AUTHORS 1 (bases 1 to 1499)  
Purthold, U., Wagner, M., Timmermann, G., Pommerening-Roseer, A. and  
Koops, H. P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
JOURNAL 13130037  
PUBMED 2 (bases 1 to 1499)  
REFERENCE Purthold, U., Timmermann, G., Koops, H. P., Pommerening-Roseer, A. and  
AUTHORS Wagner, M.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source 1. 1499  
/organism="Nitrosomonas sp. Nm47"  
/mol\_type="genomic DNA"  
/strain="Nm47"  
/db\_xref="taxon:200128"  
<1...>1499  
/product="16S ribosomal RNA"

ORIGIN  
rRNA  
Query Match 91.2%; Score 1329.4; DB 1; Length 1499;  
Best Local Similarity 94.9%; Pred. No. 7.8e-14;  
Matches 1385; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

Qy 181 GACCTTGCGCTTTTGGAGCGGCCGATGCTGATTAAGTAGTTGGTGGGTAAGGCTTAC 240  
Db 183 GACCTTAATGCTTTTGGAGCGGCCGATGCTGATTAAGTAGTTGGTGGGTAAGGCTTAC 242  
Qy 241 CAAGCGACAGATCACTAGTGTGCTTGAGAGACGACCACTGGGACTGAGACACG 300  
Db 243 CAAGCGACAGATCACTAGTGTGCTTGAGAGACGACCACTGGGACTGAGACACG 302  
Qy 301 GCCGAGCTCTTACCGGAGGACGACGATGGGGAATTTTGGACAAATGGGCGCAAGCTTGATC 360  
Db 303 GCCGAGCTCTTACCGGAGGACGACGATGGGGAATTTTGGACAAATGGGCGCAAGCTTGATC 362  
Qy 361 CAGCAATGCGCGGTGAGTGAAGAAAGCCCTGCGGTTTAAAGCTCTTCACTGAGAGAGA 420  
Db 363 CAGCAATGCGCGGTGAGTGAAGAAAGCCCTGCGGTTTAAAGCTCTTCACTGAGAGAGA 422  
Qy 421 AAAGGTTACGGTAAATTAATCGTACCAATGACGATTCGACAGAAAGACACCGGCTAAC 480  
Db 423 AAAGGTTACGGTAAATTAATCGTACCAATGACGATTCGACAGAAAGACACCGGCTAAC 482  
Qy 481 TACGTGCCAGACGCCGCGTAAATACGTAGGGTGCAGCGTTAATCGAATTAATCGGGCGT 540  
Db 483 TACGTGCCAGACGCCGCGTAAATACGTAGGGTGCAGCGTTAATCGAATTAATCGGGCGT 542  
Qy 541 AAAGGTTACGGTAAATTAATCGTACCAATGACGATTCGACAGAAAGACACCGGCTAAC 600  
Db 543 AAAGGTTACGGTAAATTAATCGTACCAATGACGATTCGACAGAAAGACACCGGCTAAC 602  
Qy 601 TCGGTTTGAATCTACCAAGCTAGAGTGGGACAGAGAGGTGGAATTCATGTGTAGCAG 660  
Db 603 TCGGTTTGAATCTACCAAGCTAGAGTGGGACAGAGAGGTGGAATTCATGTGTAGCAG 662  
Qy 661 TGAATTCGTAGAGTATGGAAGAACATCGATTCGGGAAAGGACGCTCTGGGTTTAACT 720  
Db 663 TGAATTCGTAGAGTATGGAAGAACATCGATTCGGGAAAGGACGCTCTGGGTTTAACT 722  
Qy 721 GACGCTCATGACAGAAACGTTGGGAGCAAAACAGGATTAGTACCTGGTAGTCCACGCC 780  
Db 723 GACGCTCATGACAGAAACGTTGGGAGCAAAACAGGATTAGTACCTGGTAGTCCACGCC 782  
Qy 781 CTTAAACGATGTCACTAGTGTGGGCTTTATAGCTTGTGTAACGAACCTAACCGCGTGA 840  
Db 783 CTTAAACGATGTCACTAGTGTGGGCTTTATAGCTTGTGTAACGAACCTAACCGCGTGA 842  
Qy 841 AGTTAACCGCTTGGGAGTACGTTGCGAAGTTAAATCTCAAGAAATTTGACGGGACCC 900  
Db 843 AGTTAACCGCTTGGGAGTACGTTGCGAAGTTAAATCTCAAGAAATTTGACGGGACCC 902  
Qy 901 GCACAAAGCGGTGATTTATGTGGAATTAATTCGATGCAACGCGAAAAACCTTACCTTCC 960  
Db 903 GCACAAAGCGGTGATTTATGTGGAATTAATTCGATGCAACGCGAAAAACCTTACCTTCC 962  
Qy 961 GACATGTAGCAATTTTCTAGAGATAGTATAGTCT--TCGGAAACGCTTAACACAGGTG 1017  
Db 963 GACATGTAGCAATTTTCTAGAGATAGTATAGTCT--TCGGAAACGCTTAACACAGGTG 1022  
Qy 1018 CTGCAATGCTGTGCTGACCTGCTGTCGTAAGTGTGGGTTAATGTCCTCCGCAACGAGCGCA 1077  
Db 1023 CTGCAATGCTGTGCTGACCTGCTGTCGTAAGTGTGGGTTAATGTCCTCCGCAACGAGCGCA 1082  
Qy 1078 ACCCTGTCAATTAATTTGATCATTTGTTGGGACCTTAATAGAGATGCGCGGTGACAAA 1137  
Db 1083 ACCCTGTCAATTAATTTGATCATTTGTTGGGACCTTAATAGAGATGCGCGGTGATAAA 1142  
Qy 1138 CCGAGGAAGGTGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTCACAG 1197  
Db 1143 CCGAGGAAGGTGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTCACAG 1202  
Qy 1198 TAAATCAATGGCGGTACAGAGGTTGGCAACCGCGAGGGGAGCTTAATCTCAGAAGC 1257  
Db 1203 TAAATCAATGGCGGTACAGAGGTTGGCAACCGCGAGGGGAGCTTAATCTCAGAAGC 1262



QY 1258 GCGTCGTAGTCCGATCGAGTCTGCAACTCGATCCGTAAGTCGGAATCGCTAGTAAT 1317  
|  
|  
|  
Db 1263 GCGTCGTAGTCCGATCGAGTCTGCAACTCGATCCGTAAGTCGGAATCGCTAGTAAT 1322  
|  
|  
|  
QY 1318 CCGGATCGAGATGTCGGGTGAATACGTTCCCGGGCTTGTACACCGCCCGCTACAC 1377  
|  
|  
|  
Db 1323 CCGGATCGAGATGTCGGGTGAATACGTTCCCGGGCTTGTACACCGCCCGCTACAC 1382  
|  
|  
|  
QY 1378 CATTGGAGTGGGTTTCCAGAAAGCAGTAGTCTAACCGTAAGAGGGCGCTTGCCACG 1437  
|  
|  
|  
Db 1383 CATTGGAGTGGGTTTCCAGAAAGCAGTAGTCTAACCGTAAGAGGGCGCTTGCCACG 1442  
|  
|  
|  
QY 1438 TGAGATTCACTAGCTGGGGTG 1457  
|  
|  
|  
Db 1443 TGAGATTCACTAGCTGGGGTG 1462  
|  
|  
|  
RESULT 28  
AF272422 1470 bp DNA linear BCT 05-DEC-2000  
LOCUS Nitrosomonas oligotropha 16S ribosomal RNA gene, partial sequence.  
AF272422 AF272422.1 GI:11545282  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Nitrosomonas oligotropha  
Nitrosomonas oligotropha  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
1 (bases 1 to 1470)  
Purkhof, U., Pommerening-Rosier, A., Juretschko, S., Schmid, M. C.,  
Koops, H. P., and Wagner, M.  
Phylogeny of all recognized species of ammonia oxidizers based on  
comparative 16S rRNA and amoA sequence analysis: implications for  
molecular diversity surveys  
Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)  
2 (bases 1 to 1470)  
Purkhof, U., Pommerening-Rosier, A., Juretschko, S., Schmid, M. C.,  
Koops, H. P., and Wagner, M.  
Direct Submission  
Submitted (26-May-2000) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
LOCATION/Qualifiers  
1. 1470  
/organism="Nitrosomonas oligotropha"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:42354"  
<1..>1470  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 90.7%; Score 1321.4; DB 1; Length 1470;  
Best Local Similarity 95.0%; Pred. No. 1.1e-13;  
Matches 1374; Conservative 0; Mismatches 70; Indels 3; Gaps 1;  
QY 5 AACGTGGCGGATGCTTTACATGACGAAAGCGGACACCGATGCTTGACATCG 64  
|  
|  
|  
Db 24 AACGTGGCGGATGCTTTACATGACGAAAGCGGACACCGGATGCTTGACATCG 83  
|  
|  
|  
QY 65 TGGGAGTGGCGGACGGGTGAGTAATGATCGGAACGTAATCCAGAAAGGGGGTAAACGC 124  
|  
|  
|  
Db 84 TGGGAGTGGCGGACGGGTGAGTAATGATCGGAACGTAATCCAGAAAGGGGGTAAACGC 143  
|  
|  
|  
QY 125 ATCGAAATGTGCTAATACCGCATATCTCTAGAGAGAAAGCGGGGATCGAAAGAC 184  
|  
|  
|  
Db 144 ATCGAAATGTGCTAATACCGCATATCTCTAGAGAGAAAGCGGGGATCGAAAGAC 203  
|  
|  
|  
QY 185 TTGGGCTTTGGAGGCGCGATGCTGATTAGCTAGTTGATGGGTAAAGGCTTACCAAG 244  
|  
|  
|  
Db 204 TTATGCTTTTGGAGGCGCGATGCTGATTAGCTAGTTGATGGGTAAAGGCTTACCAAG 263  
|  
|  
|  
QY 245 GCGACGATCAGTAGTTGGTCTTGAGAGAGACACGACCACTGGGACTGAGACACGGCCC 304  
|  
|  
|

Db 264 GCAACGATCAGTAGTTGGTCTGAGAGACGGCCAGCCACACTGGGACTGAGACACGGCCC 323  
|  
|  
|  
QY 305 AGACTCTTACCGGAGCGACAGTGGGAAATTTTGGCAATGGGCGCAACCTGATCCAGC 364  
|  
|  
|  
Db 324 AGACTCTTACCGGAGCGACAGTGGGAAATTTTGGCAATGGGCGCAACCTGATCCAGC 383  
|  
|  
|  
QY 365 AATGCCCGCTGAGTGAAGAAAGGCTTGGGTTGTAAAGCTCTTTCAGTGCAGAAAGAAAG 424  
|  
|  
|  
Db 384 AATGCCCGCTGAGTGAAGAAAGGCTTGGGTTGTAAAGCTCTTTCAGTGCAGAAAGAAAG 443  
|  
|  
|  
QY 425 GTTACGCTAAATATATGTAACCCATGACCGTATCGACAGAAAGACCGGCTTAATACG 484  
|  
|  
|  
Db 444 ATTGACATTAATATATGTAACCCATGACCGTATCGACAGAAAGACCGGCTTAATACG 503  
|  
|  
|  
QY 485 TGCACAGCGCGGTAATACGTAAGGTCGACGCTTAATCGGAATTAATCGGCGCTAAAG 544  
|  
|  
|  
Db 504 TGCACAGCGCGGTAATACGTAAGGTCGACGCTTAATCGGAATTAATCGGCGCTAAAG 563  
|  
|  
|  
QY 545 GGTGCCAGCGCGGCTTGTAGTCAAGTGTGAATCCCGGGCTTAACCTGGGAATTCGCG 604  
|  
|  
|  
Db 564 GGTGCCAGCGCGGCTTGTAGTCAAGTGTGAATCCCGGGCTTAACCTGGGAATTCGCG 623  
|  
|  
|  
QY 605 TTGGAACCTAACAGCTAGAGTGTGCAAGAGGATGGAATTCATGTGTACAGTGA 664  
|  
|  
|  
Db 624 TTGGAACCTAACAGCTAGAGTGTGCAAGAGGATGGAATTCATGTGTACAGTGA 683  
|  
|  
|  
QY 665 ATGCGTAGATATGGAAGAACATGATGCGGAAGCGCTCTCTGGGTTAACTGACG 724  
|  
|  
|  
Db 684 ATGCGTAGATATGGAAGAACATGATGCGGAAGCGCTCTCTGGGTTAACTGACG 743  
|  
|  
|  
QY 725 CTGATGACGAAGACCGTGGGACCAACAGATTAGTACCGTGTAGTCACGCGCTAA 784  
|  
|  
|  
Db 744 CTGATGACGAAGACCGTGGGACCAACAGATTAGTACCGTGTAGTCACGCGCTAA 803  
|  
|  
|  
QY 785 ACGATGCAACTAGTGTGGGCTTATGAGCTTGTGTAACGAAGCTAACGCGTGAAGTT 844  
|  
|  
|  
Db 804 ACGATGCAACTAGTGTGGGCTTATGAGCTTGTGTAACGAAGCTAACGCGTGAAGTT 863  
|  
|  
|  
QY 845 GACCGCTGGGAGTACGATGCGCAAGTTAAATCTCAAGGAATTCAGCGGACCGCCAC 904  
|  
|  
|  
Db 864 GACCGCTGGGAGTACGATGCGCAAGTTAAATCTCAAGGAATTCAGCGGACCGCCAC 923  
|  
|  
|  
QY 905 AAGCGGTGATATATGCGATTAATGCAATCGCAAGGAAACCTTACCTGATGACA 964  
|  
|  
|  
Db 924 AAGCGGTGATATATGCGATTAATGCAATCGCAAGGAAACCTTACCTGATGACA 983  
|  
|  
|  
QY 965 TGTAGCAATTTTCTAGAGATAGTTAGTGC--TTGCGGAACGCTTACACAGGTGCTGC 1021  
|  
|  
|  
Db 984 TGTAGCAATTTTCTAGAGATAGTTAGTGC--TTGCGGAACGCTTACACAGGTGCTGC 1043  
|  
|  
|  
QY 1022 ATGCTGTGTCAGCTGCTGTGCTGAGATGTTGGTTAAGTCCCGCAACGAGCCAAACC 1081  
|  
|  
|  
Db 1044 ATGCTGTGTCAGCTGCTGTGCTGAGATGTTGGTTAAGTCCCGCAACGAGCCAAACC 1103  
|  
|  
|  
QY 1082 TTGTCAATTAATGCGATCATTTTGGTTGGGACCTTAATGAGACTGGGGTGAACAAACCG 1141  
|  
|  
|  
Db 1104 TTGTCAATTAATGCGATCATTTTGGTTGGGACCTTAATGAGACTGGGGTGAACAAACCG 1163  
|  
|  
|  
QY 1142 AGAAGGTGGGAGATGAGTCAAGTCTCATAGCCCTTATGGGTAGGGCTTACACGTAAT 1201  
|  
|  
|  
Db 1164 AGAAGGTGGGAGATGAGTCAAGTCTCATAGCCCTTATGGGTAGGGCTTACACGTAAT 1223  
|  
|  
|  
QY 1202 ACAATGGCGGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGGCT 1261  
|  
|  
|  
Db 1224 ACAATGGCGGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGGCT 1283  
|  
|  
|  
QY 1262 CGTAGTCGCGATCGGAATGCTGCAACTGCACTCCGCTGAAGTCGGAATGCTAGTAATGGCG 1321  
|  
|  
|  
Db 1284 CGTAGTCGCGATCGGAATGCTGCAACTGCACTCCGCTGAAGTCGGAATGCTAGTAATGGCG 1343  
|  
|  
|  
QY 1322 GATCAGCATGTCGGGTGAATATGCTTCCGGGCTTGTATACACACCGCCGCTCACACCATG 1381  
|  
|  
|  
Db 1344 GATCAGCATGTCGGGTGAATATGCTTCCGGGCTTGTATACACACCGCCGCTCACACCATG 1403  
|  
|  
|

QY 1382 GGAGTGGTTTCACGAAGCAGTAGTCTAACCGTAAGAGGGCGCTTGCCAGCGGTGAG 1441  
DB 1404 TGAATGGGTTTCACGAAGCAGTAGTCTAACCGTAAGAGGGCGCTTGNNACGGGTGAG 1463  
QY 1442 ATTCAATG 1448  
DB 1464 GTTCATG 1470

RESULT 29  
AF272423 1526 bp DNA linear BCT 05-DEC-2000  
LOCUS AF272423  
DEFINITION Nitrosomonas cryocolerans 16S ribosomal RNA gene, partial sequence.  
ACCESSION AF272423  
VERSION AF272423.1 GI:11545283  
KEYWORDS  
SOURCE Nitrosomonas cryocolerans  
ORGANISM Nitrosomonas cryocolerans  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
REFERENCE  
AUTHORS 1 (bases 1 to 1526)  
Purkhold, U., Pommerening-Rose, A., Juretschko, S., Schmid, M.C.,  
Koops, H.P. and Wagner, M.  
Phylogeny of all recognized species of ammonia oxidizers based on  
comparative 16S rRNA and amoA sequence analysis: implications for  
molecular diversity surveys  
Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)  
JOURNAL  
PUBMED 11097916  
REFERENCE 2 (bases 1 to 1526)  
Purkhold, U., Pommerening-Rose, A., Juretschko, S., Schmid, M.C.,  
Koops, H.P. and Wagner, M.  
Direct Submission  
Submitted (26-MAY-2000) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1..1526  
/organism="Nitrosomonas cryocolerans"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:44575"  
<1..>1526  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 90.6%; Score 1320; DB 1; Length 1526;  
Best Local Similarity 94.9%; Pred. No. 1.1e-13;  
Matches 1386; Conservative 0; Mismatches 70; Indels 4; Gaps 2;  
QY 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAAGTCGAACGGGACGAGATGCTTGAT 60  
DB 21 ATTGAACGCTGGCGGCGATGCTTTACATGCAAGTCGAACGGGACGAGCTTGCTC 80  
QY 61 CTGGTGGCAGTGGCGGACGGGTGAGTAATGCAATCGAAAGTAATCGAAGAGGGGGTA 120  
DB 81 CTGGTGGCAGTGGCGGACGGGTGAGTAATGCAATCGAAAGTAATCGAAGAGGGGGTA 140  
QY 121 AGCGATCGAAAGATGCTTAATACCGCATATACTCTTAAGAGAGAAAGCGGGATGAAA 180  
DB 141 AGCGATCGAAAGATGCTTAATACCGCATATACTCTTAAGAGAGAAAGCGGGATGAAA 200  
QY 181 GACCTTGCGCTTTTGGAGCGGCGGATGCTGATTAGCTAGTTGGTGGGTAAAGGCTAC 240  
DB 201 GACCTTGCGCTTTTGGAGCGGCGGCGGATGCTGATTAGCTAGTTGGTGGGTAAAGGCTAC 260  
QY 241 CAAGGCGAGATCAAGTATGTTGCTTGAGAGAGACGACCACTGGGACTGAGACAG 300  
DB 261 CAAGGCGAGATCAAGTATGTTGCTTGAGAGAGACGACCACTGGGACTGAGACAG 320  
QY 301 GCCCAGACTCTTCAACGGGAGGACGAGTGGGGAATTTTGGACAATGGGCGAAAGCTGATC 360  
DB 321 GCCCAGACTCTTCAACGGGAGGACGAGTGGGGAATTTTGGACAATGGGCGAAAGCTGATC 380  
QY 361 CAGCAATGCGGGGTGAGTAAGAAAGGCTTCGGGTTGTAAGCTCTTTCAATCGAAGAA 420

DB 381 CAGCAATGCGGGGTGAGTAAGAAAGGCTTCGGGTTGTAAGCTCTTTCAATCGAAGAA 440  
QY 421 AAAGTTACGTAATTAATCTGACCATGACGATTCGACGAAGAGACCGGCTAAC 480  
DB 441 AAAGTTACGTAATTAATCTGACCATGACGATTCGACGAAGAGACCGGCTAAC 500  
QY 481 TAGGTGCCAGACCGCGGCTTAATGATAGAGGTGCAAGCGTTATCGAAATTAAGTGGCG 540  
DB 501 TAGGTGCCAGACCGCGGCTTAATGATAGAGGTGCAAGCGTTATCGAAATTAAGTGGCG 560  
QY 541 AAAGGTGCGACGCGCGCTTGTAAATCATGATGTAATATCCCGGCTTAACCTGGGAAT 600  
DB 561 AAAGGTGCGACGCGCGCTTGTAAATCATGATGTAATATCCCGGCTTAACCTGGGAAT 620  
QY 601 TCGCTTTGAATCAAAAGCTGATGCTGCGCAGAGGAGTGGAAATTCATGTGTAGCAG 660  
DB 621 TCGCTTTGAATCAAAAGCTGATGCTGCGCAGAGGAGTGGAAATTCATGTGTAGCAG 680  
QY 661 TGAATATGCTAGATATATGAGAAACATCATGAGCGAAGGACGCTCTGGGTTAACT 720  
DB 681 TGAATATGCTAGATATATGAGAAACATCATGAGCGAAGGACGCTCTGGGTTAACT 740  
QY 721 GACGCTCATGCAAGAAAGCTGGGAGCAAAACAGATTAAGATACCTGGTATCCACGCC 780  
DB 741 GACGCTCATGCAAGAAAGCTGGGAGCAAAACAGATTAAGATACCTGGTATCCACGCC 800  
QY 781 CTAAACGATGTCAACTAGTTGTTGGGCTTATAGGCTTGGTAACGAAGCTAACCGGTGA 840  
DB 801 CTAAACGATGTCAACTAGTTGTTGGGCTTATAGGCTTGGTAACCGGTGA 860  
QY 841 AGTTACCGGCTGGGAGATAGCGGTGCAAGATTAATCAAGGAATTAAGCGGGAGCC 900  
DB 861 AGTTACCGGCTGGGAGATAGCGGTGCAAGATTAATCAAGGAATTAAGCGGGAGCC 920  
QY 901 GCACAAAGCGGTGATTAATGTTGATTAATTCGATGCAAGCGGAAACCTTAACCTT 960  
DB 921 GCACAAAGCGGTGATTAATGTTGATTAATTCGATGCAAGCGGAAACCTTAACCTT 980  
QY 961 GACATGTACGAATTTTCTAGAGATGATTAAGTGC---TTCCGGAACGTTAACACAGGTG 1017  
DB 981 GACATGTACGAATTTTCTAGAGATGATTAAGTGC---TTCCGGAACGTTAACACAGGTG 1040  
QY 1018 CTGCATGCTGTCGTCAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1077  
DB 1041 CTGCATGCTGTCGTCAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1100  
QY 1078 ACCCTTGTCAATTAATGTCATCAATTTGTTGGGCACTTTAATGAGACTGCCGTGACAA 1137  
DB 1101 ACCCTTGTCAATTAATGTCATCAATTTGTTGGGCACTTTAATGAGACTGCCGTGACAA 1160  
QY 1138 CCGGAGGAAGTGGGAGATGAGCTCAAGTCTCATAGCCCTTAATGGGTAGGGCTTACACG 1197  
DB 1161 CCGGAGGAAGTGGGAGATGAGCTCAAGTCTCATAGCCCTTAATGGGTAGGGCTTACACG 1220  
QY 1198 TAAATCAATGCGGTGACAGAGGCTTGCACACCGCGAGGGGGAGCTAATCTCAGAAAGC 1257  
DB 1221 TAAATCAATGCGGTGACAGAGGCTTGCACACCGCGAGGGGGAGCTAATCTCAGAAAGC 1280  
QY 1258 GCGTGTAGTCCGATTCGAGATCTGCAACTGCACTCCGTGAAGTGGAAATGCTAGTAAT 1317  
DB 1281 GCGTGTAGTCCGATTCGAGATCTGCAACTGCACTCCGTGAAGTGGAAATGCTAGTAAT 1340  
QY 1318 CCGGATCAAGCATGTCGCGGTGAATAGTCCCGGGCTTGTATCAACACCGCCCTGACAC 1377  
DB 1341 CCGGATCAAGCATGTCGCGGTGAATAGTCCCGGGCTTGTATCAACACCGCCCTGACAC 1400  
QY 1378 CATGGAGTGGGTTTCAACGAAGCAGGATGCTTAACCGTAAGAGGGCGCTTGCACAG 1437  
DB 1401 CATGGAGTGGGTTTCAACGAAGCAGGATGCTTAACCGTAAGAGGGCGCTTGCACAG 1459  
QY 1438 TGAGATTGATGACTGGGGTG 1457

Db 1460 TGAGATTTCATGACTGGGGTG 1479

RESULT 30  
AB000699 1501 bp DNA linear BCT 06-MAR-2001  
LOCUS Nitrosomonas sp. AL212 gene for 16S rRNA, partial sequence.  
DEFINITION AB000699  
ACCESSION AB000699.1 GI:3107908  
VERSION  
KEYWORDS  
SOURCE Nitrosomonas sp. AL212  
ORGANISM Nitrosomonas sp. AL212  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
1 (eites)  
REFERENCE Suwa, Y., Sumino, T. and Moto, K.  
TITLE Phylogenetic relationships of activated sludge isolates of ammonia  
oxidizers with different sensitivities to ammonium sulfate  
J. Gen. Appl. Microbiol. 43, 373-379 (1997)  
AUTHORS Suwa, Y.  
JOURNAL Direct Submission  
TITLES Submitted (24-JAN-1997) Yuichi Suwa, National Institute for  
Resour. & Environ. Sci., Ecological Chemistry and Microbiology  
Division, 16-3 Onogawa, Tsukuba, Ibaraki 305, Japan  
REFERENCE (E-mail: suwa@nre.go.jp, Tel: +81-298-58-8318, Fax: +81-298-58-8309)  
LOCATION/Qualifiers  
FEATURES  
Source 1. 1501  
/organism="Nitrosomonas sp. AL212"  
/mol\_type="genomic DNA"  
/isolate="AL212"  
/db\_xref="taxon:153948"  
<1\_>1501  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 90.4%; Score 1317.6; DB 1; Length 1501;  
Best Local Similarity 94.4%; Pred. No. 1.2e-13;  
Matches 1378; Conservative 0; Mismatches 79; Indels 3; Gaps 1;

QY 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAATGCAACGGCGACGATGCTTCAT 60  
DB 21 ATTGAAGCTGGCGGCGATGCTTTACATGCAATGCAACGGCGACGATGCTTCAT 80  
QY 61 CTGGTGGCGATGGCGGCGATGATGATGCAATGCAATGCAATGCAATGCAATGCAAT 120  
DB 81 CCGGTGGCGATGGCGGCGATGATGATGCAATGCAATGCAATGCAATGCAATGCAAT 140  
QY 121 ACCGATGAAAGATGCTCTAATACCGATATATCTTAAGAGAAAGAGGGGATCGAAA 180  
DB 141 ACCGATGAAAGATGCTCTAATACCGATATATCTTAAGAGAAAGAGGGGATCGAAA 200  
QY 181 GACCTTCGCTTTTGAAGCGCGCGATGCTGATTAAGCTGATGGGCTTAAAGCCCTAC 240  
DB 201 GACCTTCGCTTTTGAAGCGCGCGATGCTGATTAAGCTGATGGGCTTAAAGCCCTAC 260  
QY 241 CAAAGCGACGATCTAGTATGTTGCTGAGAGGACGACGACGACGACGACGACGACGACG 300  
DB 261 CAAAGCGACGATCTAGTATGTTGCTGAGAGGACGACGACGACGACGACGACGACGACG 320  
QY 301 GCCGAGACTCTTAAGCGGAGGACGATGGGAAATTTTGACATATGGGCGCAAGCCCTATC 360  
DB 321 GCCGAGACTCTTAAGCGGAGGACGATGGGAAATTTTGACATATGGGCGCAAGCCCTATC 380  
QY 361 CAGCAATCCCGGCTGATGAGAAAGAGCCCTTGGGTTGTTAAAGCTTTTCACTCGAAGAA 420  
DB 381 CAGCAATCCCGGCTGATGAGAAAGAGCCCTTGGGTTGTTAAAGCTTTTCACTCGAAGAA 440  
QY 421 AAAGGTTACGGTAATATCGTACCGATGACGATGACGAGAAAGACCGGCTAAC 480  
DB 441 AAAGGTTACGGTAATATCGTACCGATGACGATGACGAGAAAGACCGGCTAAC 500  
QY 481 TACGTGCCAGAGCCCGGCTAATACGTAGGGTGCAAGCGTTAATCGAATTACTGGGGCT 540

Db 501 TACGTGCCAGAGCCCGGCTAATACGTAGGGTGCAAGCGTTAATCGAATTACTGGGGCT 560  
QY 541 AAAGGTTACGGGCGGCGCTTGTAAATGATGTAATATCCCGGCTTAACCTGGGAT 600  
DB 561 AAAGGTTACGGGCGGCGCTTGTAAATGATGTAATATCCCGGCTTAACCTGGGAT 620  
QY 601 TCGGTTTGAATCTAAGAGCTGATGTTGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 621 TCGGTTTGAATCTAAGAGCTGATGTTGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 680  
QY 661 TGAATGCTAGAGATGATGAGAAACATGATGCGAGGACGCTCTCGGTTTAAACACT 720  
DB 681 TGAATGCTAGAGATGATGAGAAACATGATGCGAGGACGCTCTCGGTTTAAACACT 740  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTGATACCTTGATGTCACGCC 780  
DB 741 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTGATACCTTGATGTCACGCC 800  
QY 781 CTAAACGATGTCATGATGTTGGGCTTTATTAAGCTTGTATGAGAGCTAACGGGTGA 840  
DB 801 CTAAACGATGTCATGATGTTGGGCTTTATTAAGCTTGTATGAGAGCTAACGGGTGA 860  
QY 841 AGTTGACCGGCTGGGAGGATACGGTCCGAGATTAAACTCAAGAAATTGACGGGAGCC 900  
DB 861 AGTTGACCGGCTGGGAGGATACGGTCCGAGATTAAACTCAAGAAATTGACGGGAGCC 920  
QY 901 GCACAAGCGGTGATTAATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 960  
DB 921 GCACAAGCGGTGATTAATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 980  
QY 961 GACATGATGCAATTTTCTAGATGATTAAGTCT---TCGGGAAAGCTTAACAGGGTG 1017  
DB 981 GACATGATGCAATTTTCTAGATGATTAAGTCT---TCGGGAAAGCTTAACAGGGTG 1040  
QY 1018 CTGCTAGGCTGCTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
DB 1041 CTGCTAGGCTGCTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100  
QY 1101 ACCCTGTCTAATTAATGCTGATGATTAATGATTAATGATTAATGATTAATGATTAAT 1160  
QY 1138 CCGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
DB 1161 CCGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1220  
QY 1198 TAAATCAATGCGCGTACAGAGGGTTGCCAACCGCGAGGGGAGCTTAATCTCAGAAAGC 1257  
DB 1221 TAAATCAATGCGCGTACAGAGGGTTGCCAACCGCGAGGGGAGCTTAATCTCAGAAAGC 1280  
QY 1258 GCGTCTGATGCTCGGATGCGAGTCTGCAACTCGCTCGTGAAGTCCGATCGCTAGTAAT 1317  
DB 1281 GCGTCTGATGCTCGGATGCGAGTCTGCAACTCGCTCGTGAAGTCCGATCGCTAGTAAT 1340  
QY 1318 CCGGATCAGCATGCTCGGATGCGAGTCTGCAACTCGCTCGTGAAGTCCGATCGCTAGTAAT 1377  
DB 1341 CCGGATCAGCATGCTCGGATGCGAGTCTGCAACTCGCTCGTGAAGTCCGATCGCTAGTAAT 1400  
QY 1378 CAGGAGGATGCTTCAACAGAAAGCAGTGTCTTAACCTTAAGAGAGGAGCTTCCACAGG 1437  
DB 1401 CAGGAGGATGCTTCAACAGAAAGCAGTGTCTTAACCTTAAGAGAGGAGCTTCCACAGG 1460  
QY 1438 TGAGATTCACTAGGAGTG 1457  
DB 1461 TGAGATTCACTAGGAGTG 1480

RESULT 31  
AY123794 1501 bp DNA linear BCT 23-SEP-2003  
LOCUS Nitrosomonas sp. Mml43 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY123794  
ACCESSION

VERSION AY123794.1 GI:24474423  
KEYWORDS  
SOURCE Nitrosomonas sp. Nm143  
ORGANISM Nitrosomonas sp. Nm143  
REFERENCE Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
AUTHORS Purkhold U., Wagner M., Timmermann G., Pommerening-Roesser A. and  
Koops H.P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
JOURNAL Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
PUBMED 13130037  
REFERENCE 2 (bases 1 to 1501)  
AUTHORS Purkhold U., Timmermann G., Koops H.-P., Pommerening-Roesser A. and  
Wagner M.  
TITLE Direct Submision  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
Location/Qualifiers  
1..1501  
/organism="Nitrosomonas sp. Nm143"  
/mol\_type="genomic DNA"  
/strain="Nm143"  
/db\_xref="taxon:200123"  
<1..>1501  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 90.4%; Score 1316.4; DB 1; Length 1501;  
Best Local Similarity 95.1%; Pred. No. 1.2e-13;  
Matches 1391; Conservative 0; Mismatches 66; Indels 5; Gaps 3;  
QY 1 ATTGAACGCTGGCGGCATCTTTACACATGCAAGTCAACGCGACGAGATGCTTGAT 60  
DB 3 ATTGAACGCTGGCGGCATCTTTACACATGCAAGTCAACGCGACGAGATGCTTGAT 62  
QY 61 CTGGTGGCGAGTGGCGGAGGGGTGATGATGATCGGAAGTATCCAGAAAGGGGGGTA 120  
DB 63 CTGGTGGCGAGTGGCGGAGGGGTGATGATGATCGGAAGTATCCAGAAAGGGGGGTA 122  
QY 121 ACCGATCGAAAGATGCTATATACCGCATATATCTTAAGAGAAAGACGAGGATGATA 180  
DB 123 AGCGATCGAAAGATGCTATATACCGCATATATCTTAAGAGAAAGACGAGGATGATA 182  
QY 181 GACCTTGCGCTTTTGGAGCGCGCGATGCTGATTAAGTTCAGTTGGTGAAGGCTTAC 240  
DB 183 GACCTTGCGCTTTTGGAGCGCGCGATGCTGATTAAGTTCAGTTGGTGAAGGCTTAC 242  
QY 241 CAAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 243 CAAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302  
QY 301 GCCCAGACTCTACGCGGAGGAGCAGATGGGGAATTTTGGACAATGGGCGCAAGCTGATC 360  
DB 303 GCCCAGACTCTACGCGGAGGAGCAGATGGGGAATTTTGGACAATGGGCGCAAGCTGATC 362  
QY 361 CAGCAATGCGCGTGAAGTGAAGAGGCTTTCGGGTGTAAGCTCTTTTCACTGAGAGAGA 420  
DB 363 CAGCCATGCGCGGTGAAGTGAAGAGGCTTTCGGGTGTAAGCTCTTTTCACTGAGAGAGA 422  
QY 421 AAAAGTTAGGTTAATTAATCGTGAACCATGACGGTATCGACGAAGAAACACCGGTTAC 480  
DB 423 AAAAGTTAGGTTAATTAATCGTGAACCATGACGGTATCGACGAAGAAACACCGGTTAC 482  
QY 481 TACGTGCGCAGACCGCGGTATATACGTAGGGTCAAGCTTATCGGAATTACTGGGCGT 540  
DB 483 TACGTGCGCAGACCGCGGTATATACGTAGGGTCAAGCTTATCGGAATTACTGGGCGT 542  
QY 541 AAAAGGTGCGACGCGCGCTTGTAAATGATGATGTAATCCCGGGCTTAACTTGGGAAT 600  
DB 543 AAAAGGTGCGACGCGCGCTTGTAAATGATGATGTAATCCCGGGCTTAACTTGGGAAT 602

QY 601 TGGCTTTGAAACTACAAAGCTAGAGTGGCAGAGGAGGTGAATTCATGTAGCAG 660  
DB 603 TGGCTTTGAAACTACAAAGCTAGAGTGGCAGAGGAGGTGAATTCATGTAGCAG 662  
QY 661 TGAATGCGTAGAGATATGAAAGAACATCATGATGGCAAGGACCTCTGGATTAACT 720  
DB 663 TGAATGCGTAGAGATATGAAAGAACATCATGATGGCAAGGACCTCTGGATTAACT 722  
QY 721 GACGCTATGCAACGAAGCGTGGGAGCAACAGATTAAGTACCTCTGGATTCAAGCC 780  
DB 723 GACGCTATGCAACGAAGCGTGGGAGCAACAGATTAAGTACCTCTGGATTCAAGCC 782  
QY 781 CTAAACGATGTCATGATGTTGGGCTTTTAAAGCTTGGTAAAGCTAAACGCTGA 840  
DB 783 CTAAACGATGTCATGATGTTGGGCTTTTAAAGCTTGGTAAAGCTAAACGCTGA 842  
QY 841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAATACTCAAGAGATTGACGGGACCC 900  
DB 843 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAATACTCAAGAGATTGACGGGACCC 902  
QY 901 GCACAAGCGGTGGATTAATGATGATTAATTCATGCAAGCAAGCAAGCAAGCAAGCA 960  
DB 903 GCACAAGCGGTGGATTAATGATGATTAATTCATGCAAGCAAGCAAGCAAGCAAGCA 962  
QY 961 GACATGT-AGCGAATTTTCTAGAGATAGATTAGTCC--TTCCGGGAAGCTTAACAGAGT 1016  
DB 963 GACATGTACCGAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1022  
QY 1017 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076  
DB 1023 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082  
QY 1077 AACCTTGTCAATTAATGCTCAATGCTTGGTGGGCACTTTAATGAGACTGCGCGTGA 1136  
DB 1083 AACCTTGTCAATTAATGCTCAATGCTTGGTGGGCACTTTAATGAGACTGCGCGTGA 1142  
QY 1137 ACCGAGAGAGTGGGAGTGAAGT-CAAGTCTCTAGGCGCTTATGAGTGGTACGCTTACA 1195  
DB 1143 ACCGAGAGAGTGGGAGTGAAGT-CAAGTCTCTAGGCGCTTATGAGTGGTACGCTTACA 1202  
QY 1196 GGTATATCAATGCGCGCTGACAGAGGTTGCCAACCCGAGAGGGAGCTTAATTCAGAAA 1255  
DB 1203 GGTATATCAATGCGCGCTGACAGAGGTTGCCAACCCGAGAGGGAGCTTAATTCAGAAA 1262  
QY 1256 GCGCTGCTGATGCTCGGATCGGAGTCTGCACTCGGTAAGTCCGAATTCGTTAGTA 1315  
DB 1263 GCGTGTGCTGATGCTCGGATCGGAGTCTGCACTCGGTAAGTCCGAATTCGTTAGTA 1322  
QY 1316 ATCCGAGATCAGCATGTCGCGGTGAATAGTTCCTCGGGCTTTGTAACACCGCCGCTGAC 1375  
DB 1323 ATCCGAGATCAGCATGTCGCGGTGAATAGTTCCTCGGGCTTTGTAACACCGCCGCTGAC 1382  
QY 1376 ACCATGGAGTGGGTTTACCAAGAACAGAGTATGTTAACTGTAAGAGAGGCGCTTGCAC 1435  
DB 1383 ACCATGGAGTGGGTTTACCAAGAACAGAGTATGTTAACTGTAAGAGAGGCGCTTGCAC 1442  
QY 1436 GGTGAGATTCAATGACTGGGGTG 1457  
DB 1443 GGTGAGATTCAATGACTGGGGTG 1464  
RESULT 32  
AF272414  
LOCUS AF272414 1515 bp DNA linear BCT 05-DEC-2000  
DEFINITION Nitrosomonas ureae 16S ribosomal RNA gene, partial sequence.  
ACCESSION AF272414  
VERSION AF272414.1 GI:11545274  
KEYWORDS  
SOURCE Nitrosomonas ureae  
ORGANISM Nitrosomonas ureae  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.

REFERENCE 1 (bases 1 to 1515)  
AUTHORS Purkhold,U., Pommerening-Roser,A., Juretschko,S., Schmid,M.C.,  
Koops,H.P. and Wagner,M.  
TITLE phylogeny of all recognized species of ammonia oxidizers based on  
comparative 16S rRNA and amoA sequence analysis: implications for  
molecular diversity surveys  
JOURNAL Appl. Environ. Microbiol. 66 (12), 5368-5382 (2000)  
PUBMED 11097916  
REFERENCE 2 (bases 1 to 1515)  
AUTHORS Purkhold,U., Pommerening-Roser,A., Juretschko,S., Schmid,M.C.,  
Koops,H.P. and Wagner,M.  
TITLE Direct Submission  
JOURNAL Submitted (26-MAY-2000) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
Source  
1. 1515  
/organism="Nitrosomonas ureae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:44577"  
<1..31515  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 90.3%; Score 1315.4; DB 1; Length 1515;  
Best Local Similarity 94.1%; Pred. No. 1.3e-13;  
Matches 1374; Conservative 0; Mismatches 83; Indels 3; Gaps 1;  
QY 1 ATTGAAGCTGGCGGCGATGCTTTACATGCAACTCGAAGCGGACGAGATGCTTCAT 60  
DB 2 ATTCMAAGCTGGCGGCGATGCTTTACATGCAACTCGAAGCGGACGAGATGCTTCAT 61  
QY 61 CTGGTGGCGATGGCGGACGCGGTGATGATGATCGAAGCGTATCCAGAAAGAGGGGGTA 120  
DB 62 CTGGTGGCGATGGCGGACGCGGTGATGATGATCGAAGCGTATCCAGAAAGAGGGGGTA 121  
QY 121 ACGCATGAAAGATGTGCTAATACCGCATATCTCTAAGAGAAAGCGGGATCGAAA 180  
DB 122 ACGCATGAAAGATGTGCTAATACCGCATATCTCTAAGAGAAAGCGGGATCGAAA 181  
QY 181 GACCTTGGCTTTGGAGCGGCGATGCTGATGATGATGATGATGATGATGATGATGAT 240  
DB 182 GACCTTGGCTTTGGAGCGGCGATGCTGATGATGATGATGATGATGATGATGATGAT 241  
QY 241 CAAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 242 CAAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301  
QY 301 GCCCAGATCTCTAAGCGGAGGACGATGAGGAAATTTTGAACAATGGCGCAAGCTGATC 360  
DB 302 GCCCAGATCTCTAAGCGGAGGACGATGAGGAAATTTTGAACAATGGCGCAAGCTGATC 361  
QY 361 CAGCAATGCCCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
DB 362 CAGCAATGCCCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421  
QY 421 AAAGGTTACGGTAAATATCGTACCCATGACGGTATGACAGAGAAAGACCGGCTAAC 480  
DB 422 AAAGGTTACGGTAAATATCGTACCCATGACGGTATGACAGAGAAAGACCGGCTAAC 481  
QY 481 TACGTGCGACGACCGCGGTAAATAGTAGGTGCAAGGCTTAATCGAATTAATCGAGCGT 540  
DB 482 TACGTGCGACGACCGCGGTAAATAGTAGGTGCAAGGCTTAATCGAATTAATCGAGCGT 541  
QY 541 AAAGGTTACGGTAAATATCGTACCCATGACGGTATGACAGAGAAAGACCGGCTAAC 600  
DB 542 AAAGGTTACGGTAAATATCGTACCCATGACGGTATGACAGAGAAAGACCGGCTAAC 601  
QY 601 TGGGTTTAAACTACAAAGCTAGATGATGATGATGATGATGATGATGATGATGATGAT 660  
DB 602 TGGGTTTAAACTACAAAGCTAGATGATGATGATGATGATGATGATGATGATGATGAT 661  
QY 661 TGAATGCGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

DB 662 TGAATGCGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721  
QY 721 GACGCTATGACGAAAGCTGGGGAGCAACAAGATTAGTACCTGGTATGACCGC 780  
DB 722 GACACTCAGGACGAAAGCTGGGGAGCAACAAGATTAGTACCTGGTATGACCGC 781  
QY 781 CTAAAGATGATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 782 CTAAAGATGATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 841  
QY 841 AGTTGACCGCTGGGGAGTACCGTTCGAAGTTAAATCTCAAGAAATGACGGGACCC 900  
DB 842 AGTTGACCGCTGGGGAGTACCGTTCGAAGTTAAATCTCAAGAAATGACGGGACCC 901  
QY 901 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 902 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961  
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017  
DB 962 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021  
QY 1018 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
DB 1022 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081  
QY 1078 ACCCTGTCTAATTAATGCTATCTATGTTGGGCACTTATATGATGATGATGATGAT 1137  
DB 1082 ACCCTGTCTAATTAATGCTATCTATGTTGGGCACTTATATGATGATGATGATGAT 1141  
QY 1138 CCGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
DB 1142 CCGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201  
QY 1198 TAATACAAATGGCGGTACAGAGGATGCTCAACCGCGAGGGGAGCTTAATCTCAAAAGC 1257  
DB 1202 TAATACAAATGGCGGTACAGAGGATGCTCAACCGCGAGGGGAGCTTAATCTCAAAAGC 1261  
QY 1258 GCGTGTATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317  
DB 1262 GCGTGTATGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1321  
QY 1318 CCGGATCAGCATGCTCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377  
DB 1322 CCGGATCAGCATGCTCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381  
QY 1378 CATGGAGTGGGTTTACAGAGAGTATGATGATGATGATGATGATGATGATGATGAT 1437  
DB 1382 CATGGAGTGGGTTTACAGAGAGTATGATGATGATGATGATGATGATGATGATGAT 1441  
QY 1438 TGAGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
DB 1442 TGAGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1461  
RESULT 33  
AJ621026 1454 bp DNA linear BCT 08-JUN-2004  
LOCUS Nitrosomonas sp. 1679A3 16S rRNA gene, isolate 1679A3.  
DEFINITION ACCESSION  
AJ621026  
VERSION AJ621026.1 GI:40994845  
KEYWORDS 16S ribosomal RNA; 16S rRNA gene.  
SOURCE Nitrosomonas sp. 1679A3  
ORGANISM Nitrosomonas sp. 1679A3  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosomonas.  
REFERENCE 1  
AUTHORS Nielsen,S., Baer Gillisen,M.J., Revsbech,N., Laanbroek,H.J. and  
Boilmann,A.  
TITLE Influence of heterotrophic bacteria on the activity and growth  
behavior of ammonia-oxidizing bacteria  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1454)

AUTHORS Bollmann, A.  
TITLE Direct Subdivision  
JOURNAL Submitted (15-JAN-2004) Bollmann A., Department for Microbial  
Ecology, NIOO-KNAW Centre for Limnology, Rijksstraatweg 6, 3631 AC  
Nieuwerluis, NETHERLANDS  
FEATURES  
source 1. .1454  
/organism="Nitrosomonas sp. Is79A3"  
/mol\_type="genomic DNA"  
/isolate="Is79A3"  
/isolation\_source="fresh water sediment"  
/db\_xref="taxon:261292"  
/country="Netherlands:lake Drontemeer"  
gene 1. .1449  
/gene="16S rRNA"  
rRNA 1. .1449  
/gene="16S rRNA"  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 90.0%; Score 1310.6; DB 1; Length 1454;  
Best Local Similarity 95.3%; Pred. No. 1.6e-13;  
Matches 1384; Conservative 0; Mismatches 64; Indels 5; Gaps 3;  
QY 1 ATTGAACGCTGGCGCATGCTTTACATGCAAGTGGAAACGACGACGAGATGCTGCAT 60  
DB 4 ATTGAACGCTGGCGCATGCTTTACATGCAAGTGGAAACGACGACGAGATGCTGCAC 63  
QY 61 CTGGTGGCGATGGCGGACGGGTGAGTAATGCATCGGAAAGTATCCAGAAAGAGGGGGTA 120  
DB 64 CTGGTGGCGATGGCGGACGGGTGAGTAATGCATCGGAAAGTATCCAGAAAGTATCCAG 122  
QY 121 AGCGATCGAAAGATGCTAATACCGCATATCTCTAAGAGGAAAGCGGGGATGAGAA 180  
DB 123 AGCGATCGAAAGATGCTAATACCGCATATCTCTAAGAGGAAAGTATCCAGAAAGTAT 182  
QY 181 GACCTTGCGCTTTTGGAGCGCGCATGCTGATTAGCTGGGTGAAAGGCTTAC 240  
DB 183 GACCTTGCGCTTTTGGAGCGCGCATGCTGATTAGCTGGGTGAAAGGCTTAC 242  
QY 241 CAAGCGACGATCAGTAGTTGGTCTTGAGAGAGCAGACCAAGCTGGGACTGAGACAG 300  
DB 243 CAAGCGACGATCAGTAGTTGGTCTTGAGAGAGCAGACCAAGCTGGGACTGAGACAG 302  
QY 301 GCCCAGACTCTTACCGGAGGACAGAGTGGGAAATTTTGAACAATGSGGCCAAGCTGATC 360  
DB 303 GCCCAGACTCTTACCGGAGGACAGAGTGGGAAATTTTGAACAATGSGGCCAAGCTGATC 362  
QY 361 CAGCAATGCGCGTGAAGTGAAGAGGCTTCGGGTGTAAGCTTTTCATGTCAGAGAA 420  
DB 363 CAGCAATGCGCGTGAAGTGAAGAGGCTTCGGGTGTAAGCTTTTCATGTCAGAGAA 422  
QY 421 AAAGGTTACGTTAATATCTGACCCATGACGGTATCGACAGAAAGACACCGGCTAAC 480  
DB 423 AAAGGTTGACATTAATCATATCATGACGGTATCGACAGAAAGACACCGGCTAAC 482  
QY 481 TACGTGCCAGACCGCGGTAACTAGTGGTGCAAGCCTTAATCGGAATTTCTGGGCT 540  
DB 483 TACGTGCCAGACCGCGGTAACTAGTGGTGCGAGCCTTAATCGGAATTTCTGGGCT 542  
QY 541 AAAGGTTGCGACGCGCTTGTAAATGATGATGTGAATCCCGGGCTTAACCTGGGAAT 600  
DB 543 AAAGGTTGCGACGCGCTTGTAAATGATGATGTGAATCCCGGGCTTAACCTGGGAAT 602  
QY 601 TCGGTTTGAACCTAACAAGCTAAGTGTGGCAGAGGAGTGGAAATTCATGTGTAGCAG 660  
DB 603 TCGGTTTGAACCTAACAAGCTAAGTGTGGCAGAGGAGTGGAAATTCATGTGTAGCAG 662  
QY 661 TGAATATGCTAAGATATGGAAGAACTAGATGCGCAAGGACCTCCCTGGGTTAACT 720  
DB 663 TGAATATGCTAAGATATGGAAGAACTAGATGCGCAAGGACCTCCCTGGGTTAACT 722  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGTGTAGTCCAGCGC 780

DB 723 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGTGTAGTCCAGCGC 782  
QY 781 CTAAACGATGTCACTAGTTGTTGGGCTTATTAGCTTTGGTAAGAACTAACCGGTGA 840  
DB 783 CTAAACGATGTCACTAGTTGTTGGGCTTATTAGCTTTGGTAAGAACTAACCGGTGA 842  
QY 841 AGTTGACCGCTGGGAGATGACGGTGCAGAAATTAACCTCAAGAAATGACGGGAGCC 900  
DB 843 AGTTGACCGCTGGGAGATGACGGTGCAGAAATTAACCTCAAGAAATGACGGGAGCC 902  
QY 901 GCACAAGCGGTGATTAATGTGGATTAATTCGATGCAACGCGAAAACTTACCTACCTT 960  
DB 903 GCACAAGCGGTGATTAATGTGGATTAATTCGATGCAACGCGAAAACTTACCTACCTT 962  
QY 961 GACATGTACCGAATTTTCTAGAGATAGATTAGTGTCT--TGGGAACGCTAACACAGGTG 1017  
DB 963 GACATGTACCGAATTTTCTAGAGATAGATTAGTGTCTCGAAGAAATTTGAACACAGGTG 1022  
QY 1018 CTGCATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
DB 1023 CTGCATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082  
QY 1078 ACCCTTGCTAATTAATGCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
DB 1083 ACCCTTGCTAATTAATGCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142  
QY 1138 CCGAGAGAAAGTGGGAGATGACGTCAGATCTCATATGCCCCCTTAATGGGTAGGCTTACACG 1197  
DB 1143 CCGAGAGAAAGTGGGAGATGACGTCAGATCTCATATGCCCCCTTAATGGGTAGGCTTACACG 1202  
QY 1198 TAAATCAATGGGCGCTACAGAGAGGTTGCAACCCGAGAGGGAGCTAATCTCAGAAAGC 1257  
DB 1203 TAAATCAATGGGCGCTACAGAGAGGTTGCAACCCGAGAGGGAGCTAATCTCAGAAAGC 1262  
QY 1258 GCGTGTAGTCCGATTCGAGATCTGCAACCTGACTCGGTGAAGTGGAAATGCTAGTAAT 1317  
DB 1263 GCGTGTAGTCCGATTCGAGATCTGCAACCTGACTCGGTGAAGTGGAAATGCTAGTAAT 1322  
QY 1318 CCGGATCAGCATGTCGGGTGTAATAGTTCGCCGGGTCTTGTACACACCCCGCTCACAC 1377  
DB 1323 CCGGATCAGCATGTCGGGTGTAATAGTTCGCCGGGTCTTGTACACACCCCGCTCACAC 1382  
QY 1378 CATGGAGTGGGTTTACACAGAAAGGATGCTAACCGTAAAGAGGCGCTTGGCACAG 1437  
DB 1383 CATGGAGTGGGTTTACACAGAAAGGATGCTAACCGTAAAGAGGCGCTTGGCACAG 1441  
QY 1438 TGAGATTCATGAC 1450  
DB 1442 TGAGATTCATGAC 1454  
RESULT 34  
AY123806  
LOCUS 1497 bp DNA linear BCT 23-SEP-2003  
DEFINITION Nitrosospiira sp. Kaz 16S ribosomal RNA gene, partial sequence.  
ACCESSION AY123806  
VERSION AY123806.1 GI:24474435  
KEYWORDS  
SOURCE  
ORGANISM Nitrosospiira sp. Kaz  
Nitrosospiira sp. Kaz  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospiira.  
REFERENCE  
1 (bases 1 to 1497)  
Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roesser, A. and  
Koops, H.-P.  
16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
JOURNAL  
PUBMED 13130037  
REFERENCE  
2 (bases 1 to 1497)  
Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roesser, A. and





/organism="Nitrosospirita sp. Nsp57"  
/mol\_type="genomic DNA"  
/strain="Nsp57"  
/db\_xref="taxon:200118"  
<1..>1497  
/product="16S ribosomal RNA"

## ORIGIN

```
Query Match      89.4%; Score 1302; DB 1; Length 1497;
Best Local Similarity 94.2%; Pred. No. 2.1e-33;
Matches 1375; Conservative 0; Mismatches 80; Indels 5; Gaps 2;

QY 1 ATGAAAGCTGGGGGATGCTTACACATGCAAGTGAACGGAGACGGATCTTGCAAT 60
   |||
DB 3 ATGAAAGCTGGGGGATGCTTACACATGCAAGTGAACGGAGACGGATCTTGCAAT 60
   |||

QY 61 CTGGTGGCGAGTGGCGGACGGGTGAGTATGATCGGAACGTATCCAGAGAGGGGGAT 120
   |||
DB 61 CTGGTGGCGAGTGGCGGACGGGTGAGTATGATCGGAACGTATCCAGAGAGGGGGAT 120
   |||

QY 121 AGCGATCGAAAGTGTGCTAATACCGCATATCTCTTAAAGAGAAAGCAGGGATCGAAA 180
   |||
DB 121 AGCGATCGAAAGTGTGCTAATACCGCATATCTCTTAAAGAGAAAGCAGGGATCGAAA 180
   |||

QY 181 GACCTTGGCGCTTTTGAAGGGCGCGATGCTGATTAAGTGTGGGGGTTAAAGGCTTAC 240
   |||
DB 181 GACCTTGGCGCTTTTGAAGGGCGCGATGCTGATTAAGTGTGGGGGTTAAAGGCTTAC 240
   |||

QY 241 CAAGCGAGATCATGATGTTGTCTGAGAGAGCAGACAGCAGCAGCTGGAGCTGAGACAG 300
   |||
DB 241 CAAGCGAGATCATGATGTTGTCTGAGAGAGCAGACAGCAGCAGCTGGAGCTGAGACAG 300
   |||

QY 301 GCCCAGACTCTCTACGGGAGGACAGACAGTGGGGAAATTTTGGCAATGGCGCGACCTGATC 360
   |||
DB 301 GCCCAGACTCTCTACGGGAGGACAGACAGTGGGGAAATTTTGGCAATGGCGCGACCTGATC 360
   |||

QY 361 CAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGGAGAG 420
   |||
DB 361 CAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGGAGAG 420
   |||

QY 421 AAAGTTACGTTAAATATATCTGTGACCCATGACCGGTATCGACAGAGAGACCGGCTTAA 480
   |||
DB 421 AAAGTTACGTTAAATATATCTGTGACCCATGACCGGTATCGACAGAGAGACCGGCTTAA 480
   |||

QY 481 TACGTGCCAGACCGCGGTAAATAGTAAAGGTGCAAGCGTTATTCGGAATTAATCTGGCGGT 540
   |||
DB 481 TACGTGCCAGACCGCGGTAAATAGTAAAGGTGCAAGCGTTATTCGGAATTAATCTGGCGGT 540
   |||

QY 541 AAAGGTTGGGAGCGGCGCTTGTAAAGTCAAGATGTAATCCCGGGCTTAACTTGGGAAT 600
   |||
DB 541 AAAGGTTGGGAGCGGCGCTTGTAAAGTCAAGATGTAATCCCGGGCTTAACTTGGGAAT 600
   |||

QY 601 TGGGTTGAAACTCAAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660
   |||
DB 601 TGGGTTGAAACTCAAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660
   |||

QY 661 TGAATATGCGTAAAGATATGAGAGAAACATGATGGCGAGCGACCTCTGGGTTAACTACT 720
   |||
DB 661 TGAATATGCGTAAAGATATGAGAGAAACATGATGGCGAGCGACCTCTGGGTTAACTACT 720
   |||

QY 721 GACGCTCATGACGAAGAGCGTGGGAGCAAAACAGATTAGATACCTCTGGTATGCCAGCC 780
   |||
DB 721 GACGCTCATGACGAAGAGCGTGGGAGCAAAACAGATTAGATACCTCTGGTATGCCAGCC 780
   |||

QY 781 CTAAACGATGTCAACTAGTTGTGGGCTTATTAAGGCTTGTATACGACGTAACGCGTGA 840
   |||
DB 781 CTAAACGATGTCAACTAGTTGTGGGCTTATTAAGGCTTGTATACGACGTAACGCGTGA 840
   |||

QY 841 AGTTGACCGCTGGGGAGTACGCTCGCAAGATTTAAACCTCAAGGAATTTGACGGGGACCC 900
   |||
DB 841 AGTTGACCGCTGGGGAGTACGCTCGCAAGATTTAAACCTCAAGGAATTTGACGGGGACCC 900
   |||

QY 901 GCACAAAGCGGTGATTAATGTGATTAATTCGATGCAACGCAAAACCTTACTACTGCTT 960
   |||
```

```
DB 901 GCACAAAGCGGTGATTAATGTGATTAATTCGATGCAACGCAAAACCTTACTACTGCTT 960
   |||

QY 961 GACATGTACGGAATTTTCTAGAGATGATTAGTGC---TTGGGAAACGTTAACACAGGTG 1017
   |||
DB 961 GACATGTACGGAATTTTCTAGAGATGATTAGTGC---TTGGGAAACGTTAACACAGGTG 1020
   |||

QY 1018 CTGCATGGCTGTCGTCAGCTGTCGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1077
   |||
DB 1021 CTGCATGGCTGTCGTCAGCTGTCGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1080
   |||

QY 1078 ACCCTTGTCAATTAATTTGCAATATTTGGTTGGGCACTTAAATGAGACTGCCGCTGACAA 1137
   |||
DB 1081 ACCCTTGTCAATTAATTTGCAATATTTGAGTTGGGCACTTAAATGAGAACTGCCGCTGACAA 1140
   |||

QY 1138 CCGAGAGAAAGTGGGAGTACGTCAGATCTCATGCGCTTATGGGTAGGGCTTACACAG 1197
   |||
DB 1141 CCGAGAGAAAGTGGGAGTACGTCAGATCTCATGCGCTTATGGGTAGGGCTTACACAG 1200
   |||

QY 1198 TAAATCAATGGCGGTACAGAGGCTTGCACACCGCGAGGGGAGCTAATCTCAGAAAGC 1257
   |||
DB 1201 TAAATCAATGGCGGTACAGAGGCTTGCACACCGCGAGGGGAGCTAATCTCAGAAAGC 1260
   |||

QY 1258 GCGTGTAGTCCGATTCGAGTCTGCAACTCGACTCCGTGAAGTGGAAATGCTAGTAAT 1317
   |||
DB 1261 GCGTGTAGTCCGATTCGAGTCTGCAACTCGACTCCGTGAAGTGGAAATGCTAGTAAT 1320
   |||

QY 1318 CGCGATCATGATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGCTCACAC 1377
   |||
DB 1321 CGCGATCATGATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGCTCACAC 1380
   |||

QY 1378 CATGGAGTGGGTTTCCACAGAGAGGATGCTAACCGTAAAGAGAGGGCGCTTGGCACAG 1437
   |||
DB 1381 CATGGAGTGGGTTTCCACAGAGAGGATGCTAACCGTAAAGAGAGGGCGCTTGGCACAG 1440
   |||

QY 1438 TGAGATTCAATGACTGGGGTG 1457
   |||
DB 1441 TGAGATTCAATGACTGGGGTG 1460
   |||

RESULT 36
AY123809 1497 bp DNA linear BCT 23-SEP-2003
LOCUS Nitrosospirita sp. II17 16S ribosomal RNA gene, partial sequence.
DEFINITION AY123809.1
ACCESSION AY123809.1 GI:24474438
VERSION
KEYWORDS
SOURCE
ORGANISM Nitrosospirita sp. II17
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosospirita.
REFERENCE
1 (bases 1 to 1497)
Purthold,U., Wagner,M., Timmermann,G., Pommerening-Roseer,A. and
Koops,H.P.
16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial
ammonia-oxidizing isolates: extension of the dataset and proposal
of a new lineage within the nitrosomonads
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)
JOURNAL
PUBMED 13130037
REFERENCE
2 (bases 1 to 1497)
Purthold,U., Timmermann,G., Koops,H.-P., Pommerening-Roseer,A. and
Wagner,M.
Direct Submission
Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES
source
1..1497
/organism="Nitrosospirita sp. II17"
/mol_type="genomic DNA"
/strain="II17"
/db_xref="taxon:174932"
<1..>1497
/product="16S ribosomal RNA"
```



## ORIGIN

Query Match 89.4%; Score 1302; DB 1; Length 1497;  
 Best Local Similarity 94.2%; Pred. No. 2, 1e-13;  
 Matches 1375; Conservative 0; Mismatches 80; Indels 5; Gaps 2;

1 ATTGAACGCTGGCGGAGTCTTTTACATGCAAGTGCAGCGGAGCAGACGGAGCTTGCAT 60  
 3 ATTGAACGCTGGCGGAGTCTTTTACATGCAAGTGCAGCGGAGCAGACGGAGCTTGCAT 60  
 61 CTGGTGGCGAGTGGCCGAGCGGGTGAATGCAATCGGAACGTATCCAGAAAGAGGGGGGTA 120  
 61 CTGGTGGCGAGTGGCCGAGCGGGTGAATGCAATCGGAACGTATCCAGAAAGAGGGGGGTA 120  
 121 AGGCATCGAAAGATGTCTAATACCGCATATATCTTAAGAGGAAAGAGCGGGATCGAAA 180  
 121 AGGCATCGAAAGATGTCTAATACCGCATATATCTTAAGAGGAAAGAGCGGGATCGAAA 180  
 181 GACCTTGCCTTTTGGAGCGGCGCATGTCTGAATTAAGTGTGGTGGGTTAAAGCCCTAC 240  
 181 GACCTTGCCTTTTGGAGCGGCGCATGTCTGAATTAAGTGTGGTGGGTTAAAGCCCTAC 240  
 241 CAAGCGACGATCATGTAATGTTGTTCTGAGAGAGCAACGACCACTGGGACTGAGACAG 300  
 241 CAAGCGCTTCATGATGCTGTCTGAGAGAGCAACGACCACTGGGACTGAGACAG 300  
 301 GCCCAACTCTCTACCGGAGGACAGTGGGGAAATTTTGAACAATGGGCGGAGCTGATC 360  
 301 GCCCAACTCTCTACCGGAGGACAGTGGGGAAATTTTGAACAATGGGCGGAGCTGATC 360  
 361 CAGCAATGCGCGCTGAGTGAAGAAAGCCCTTCGGGTTGTAAGCTCTTTCAGTGGAGAA 420  
 361 CAGCAATGCGCGCTGAGTGAAGAAAGCCCTTCGGGTTGTAAGCTCTTTCAGTGGAGAA 420  
 421 AAAAGTTACGTTAAATATGTGACCCATGACGCTATGCAAGAAAGACACCGGCTTAC 480  
 421 AAAAGTTACGTTAAATATGTGACCCATGACGCTATGCAAGAAAGACACCGGCTTAC 480  
 481 TAGCTGCACAGCGCGGGTAAATGCGTAGGGTGCAGCGTTAATCGGAATTTACTGGGCGT 540  
 481 TAGCTGCACAGCGCGGGTAAATGCGTAGGGTGCAGCGTTAATCGGAATTTACTGGGCGT 540  
 481 TAGCTGCACAGCGCGGGTAAATGCGTAGGGTGCAGCGTTAATCGGAATTTACTGGGCGT 540  
 541 AAAAGGTGCGCAGCGCGCTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTGGGAAT 600  
 541 AAAAGGTGCGCAGCGCGCTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTGGGAAT 600  
 601 TGCCTTTGAAACTTCAAAAGCTAAGTGTGCGAGAGGAGTGGAAATTCATGTGTACAG 660  
 601 TGCCTTTGAAACTTCAAAAGCTAAGTGTGCGAGAGGAGTGGAAATTCATGTGTACAG 660  
 661 TGAATATCGCTAGATATGGAAGAACATCGATGCGGAAGGACGCTCGGGTTAACTACT 720  
 661 TGAATATCGCTAGATATGGAAGAACATCGATGCGGAAGGACGCTCGGGTTAACTACT 720  
 721 GAGCCTCATGCAAGAAAGCTGGGAGAGCAACAGATTAAGTACCTGTAGTACAGCC 780  
 721 GAGCCTCATGCAAGAAAGCTGGGAGAGCAACAGATTAAGTACCTGTAGTACAGCC 780  
 781 CTAAACGATGCACTAGTGTGGGCTTAAATTAAGCTTGTGAACGAAGCTAACGCGTGA 840  
 781 CTAAACGATGCACTAGTGTGGGCTTAAATTAAGCTTGTGAACGAAGCTAACGCGTGA 840  
 841 AGTTGACCGCTGGGAGTACGGTGCAGAAATTAACGAAAGAAATTAACGCGGAGACC 900  
 841 AGTTGACCGCTGGGAGTACGGTGCAGAAATTAACGAAAGAAATTAACGCGGAGACC 900  
 901 GCACAAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTT 960  
 901 GCACAAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTT 960  
 961 GACATGTAGCGAATTTTCTAGAGATTAATTAAGTGC--TTGGGAAAGCTTAACACAGGTG 1017  
 961 GACATGTAGCGAATTTTCTAGAGATTAATTAAGTGC--TTGGGAAAGCTTAACACAGGTG 1017  
 1020 GACATGTAGCGAATTTTCTAGAGATTAATTAAGTGC--TTGGGAAAGCTTAACACAGGTG 1020

1018 CTGCATGCGCTGTCTGACGCTGCTGTGCTGAGATGTTGGGTTAAAGTCCGCAACGAGCGCA 1077  
 1021 CTGCATGCGCTGTCTGACGCTGCTGTGCTGAGATGTTGGGTTAAAGTCCGCAACGAGCGCA 1080  
 1078 ACCCTTGTCAATTAATGTCATCATTTGGTGGGCACTTTAATGAGACTGCGGTGACAAA 1137  
 1081 ACCCTTGTCAATTAATGTCATCATTTAGTGGGCACTTTAATGAAATGCGCGTGAACAA 1140  
 1138 CCGAGAAAGTGGGAGATGACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTCAACAG 1197  
 1141 CCGAGAAAGTGGGAGATGACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTCAACAG 1200  
 1198 TAATACATAGGCGCTCAACGAGGTTCCAAACCGGAGGGGAGGATTAATCAAGAAC 1257  
 1201 TAATACATAGGCGCTCAACGAGGTTCCAAACCGGAGGGGAGGATTAATCAAGAAC 1260  
 1258 GGGTGTAGTCCGAGATCGGAGTCTGCAACTCGACTCCGTGAAGTGGAAATCGCTAAT 1317  
 1261 GGGTGTAGTCCGAGATCGGAGTCTGCAACTCGACTCCGTGAAGTGGAAATCGCTAAT 1320  
 1318 CCGGATCAGCATGTCTGCGGTGAATACGTTCCGGGCTTTGTACACACCGCCGCTACAC 1377  
 1321 CCGGATCAGCATGTCTGCGGTGAATACGTTCCGGGCTTTGTACACACCGCCGCTACAC 1380  
 1378 CATGGAGTGGGTTTCAACGAAGCAGTATCTTAACCGTAAGAGGGGCGCTTGCACAG 1437  
 1381 CATGGAGTGGGTTTCAACGAAGCAGTATCTTAACCGTAAGAGGGGCGCTTGCACAG 1440  
 1438 TGAATTCATGACTGGGGTG 1457  
 1441 TGAATTCATGACTGGGGTG 1460

RESULT 37  
 AY123789 1497 bp DNA linear BCT 23-SEP-2003  
 LOCUS Nitrosospira sp. Nsp58 16S ribosomal RNA gene, partial sequence.  
 DEFINITION AY123789  
 ACCESSION AY123789  
 VERSION AY123789.1 GI:24474418  
 KEYWORDS Nitrosospira sp. Nsp58  
 SOURCE Nitrosospira sp. Nsp58  
 ORGANISM Nitrosomonadaceae; Betaproteobacteria; Nitrosomonadales; Bacteria; Proteobacteria; Nitrosospira.  
 REFERENCE 1 (bases 1 to 1497)  
 Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roser,A. and Koops,H.P.  
 TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal of a new lineage within the nitrosomonade  
 Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
 JOURNAL 13130037  
 PUBMED 2 (bases 1 to 1497)  
 REFERENCE Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Roeser,A. and Wagner,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-JUN-2002) Leibniz-Institut fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
 FEATURES  
 source  
 1. .1497  
 /organism="Nitrosospira sp. Nsp58"  
 /mol\_type="genomic DNA"  
 /strain="Nsp58"  
 /db\_xref="taxon:200116"  
 <1..>1497  
 /product="16S ribosomal RNA"

## ORIGIN

Query Match 89.3%; Score 1300.4; DB 1; Length 1497;  
 Best Local Similarity 94.1%; Pred. No. 2, 2e-13;  
 Matches 1374; Conservative 0; Mismatches 81; Indels 5; Gaps 2;

QY	1	ATTGAACGCTGGCGGCAATGCTTTACATATGCAAGTCGAA	CGGCAAGACGGAATGCTTGCAAT	60
Db	3	ATTGAACGCTGGCGGCAATGCTTTACATATGCAAGTCGAA	CGGCAAGACGGAATGCTTGCAAT	60
QY	61	CTGGTGGCGAATGGCGGAGCGGGTAGTAATGCAATCGAA	CGTATCCAGAAAGAGGGGGGATA	120
Db	61	CTGGTGGCGAATGGCGGAGCGGGTAGTAATGCAATCGAA	CGTATCCAGAAAGAGGGGGGATA	120
QY	121	ACGCATCGAAGAATGTGCTAATACGCATATATCTTAAAGAGAAAGCAGGGGATCGAAA		180
Db	121	ACGCATCGAAGAATGTGCTAATACGCATATATCTTAAAGAGAAAGCAGGGGATCGAAA		180
QY	181	GACCTTCGGCTTTGGAGCGGCGCGATGTCTGAATAGCTAATGTTGGTGGGTTAAAGCTTAC		240
Db	181	GACCTTCGGCTTTGGAGCGGCGCGATGTCTGAATAGCTAATGTTGGTGGGTTAAAGCTTAC		240
QY	241	CAAGGCGACGATCAGTAATGGTGGCTGAGAGAGAGACACGACCCACATGGGGATCGAGACAG		300
Db	241	CAAGGCGATCAGTAATGGTGGCTGAGAGAGAGACACGACCCACATGGGGATCGAGACAG		300
QY	301	GCCCAAGACTCTTACCGGAGAGCGACAGTGGGGAAATTTGGACATGGGCGCAAGCCTGATC		360
Db	301	GCCCAAGACTCTTACCGGAGAGCGACAGTGGGGAAATTTGGACATGGGCGCAAGCCTGATC		360
QY	361	CAGCAATGCGCGGTGATGTAAGAGAGCCCTTCGGGTTGTAAAGCTCTTTCAATGCGAAGAGA		420
Db	361	CAGCAATGCGCGGTGATGTAAGAGAGCCCTTCGGGTTGTAAAGCTCTTTCAATGCGAAGAGA		420
QY	421	AAAGGTTACGGTAATTAATCGTAGCCCATGACGGTATCGACAGAAAGAACACGGGCTAAC		480
Db	421	AAAGGTTACGGTAATTAATCGTAGCCCATGACGGTATCGACAGAAAGAACACGGGCTAAC		480
QY	481	TACGTGCCAGACAGCCGCGTAAATACGTAGGAGTGCAAGCCTTAATCGAATTACTGGCGGT		540
Db	481	TACGTGCCAGACAGCCGCGTAAATACGTAGGAGTGCAAGCCTTAATCGAATTACTGGCGGT		540
QY	541	AAAGGTTGGCGAGCGCGGCTTTGAATGTCAGATGTGAAATCCCGGGCTTTAACTGGGAAT		600
Db	541	AAAGGTTGGCGAGCGCGGCTTTGAATGTCAGATGTGAAATCCCGGGCTTTAACTGGGAAT		600
QY	601	TGCGTTTAAACTTAAGAGCTAGATGTGGGCGGAGAGAGAGTGGAAATTCATGTGTAGACAG		660
Db	601	TGCGTTTAAACTTAAGAGCTAGATGTGGGCGGAGAGAGAGTGGAAATTCATGTGTAGACAG		660
QY	661	TGAATGCGTAGAGATATATGAAAGAAATCATGATGGCGAAGGCGAGCCTCTGGGTTAACT		720
Db	661	TGAATGCGTAGAGATATATGAAAGAAATCATGATGGCGAAGGCGAGCCTCTGGGTTAACT		720
QY	721	GACGCTCATGACGAAAGCGTGGGGAGCAAAACAGAAATTAATACCTTGGTAGTCCACGCG		780
Db	721	GACGCTCATGACGAAAGCGTGGGGAGCAAAACAGAAATTAATACCTTGGTAGTCCACGCG		780
QY	781	CTAAACGATGTCAACTAGTGTGTGGGCGCTTAATTAAGCTTGGTGTAAAGAACTTAACGCGTGA		840
Db	781	CTAAACGATGTCAACTAGTGTGTGGGCGCTTAATTAAGCTTGGTGTAAAGAACTTAACGCGTGA		840
QY	841	AGTTACCGGCTGGGGAGTAAGCGGTGCGAAGATTAAATCTCAAGAGAAATTGACGGGGACCC		900
Db	841	AGTTACCGGCTGGGGAGTAAGCGGTGCGAAGATTAAATCTCAAGAGAAATTGACGGGGACCC		900
QY	901	GCACAAAGCGGTGAGATTATGTGGATTAAATTCGATGACGACGGCAAAACCTTAACCTTACCTT		960
Db	901	GCACAAAGCGGTGAGATTATGTGGATTAAATTCGATGACGACGGCAAAACCTTAACCTTACCTT		960
QY	961	GACATGTAGCGAAATTTTCTAAGAGATAGATTAATGTC---TTGGGAAAGCTTAACAGAGTG		1017
Db	961	GACATGTAGCGAAAGCCCGCGAGAGAGCGGGCGTGGCCGAAAGAGAAACGTATAACAGAGTG		1020
QY	1018	CTGCATGGCTGTGCGTCAAGCTGTGCGTGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA		1077
Db	1021	CTGCATGGCTGTGCGTCAAGCTGTGCGTGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA		1080
QY	1078	ACCCTGTGATTAATTTGCAATCATTTTGGTTGGGCACTTTAATGACATGCGCGGTGACAA		1137

Db	1061	ACCTTGTGATTAAATGCGATTCATTCAGTGGGACCTTTAATGAACACTGCCGGTGACMAA	1140
QY	1138	CCGGAGGAAAGGTGGGGATGACGTCMAGTCTCATGCGCCTTATGGGTAGGGCTTCAACG	1197
Db	1141	CCGGAGGAAAGGTGGGGATGACGTCMAGTCTCATGCGCCTTATGGGTAGGGCTTCAACG	1200
QY	1198	TAAATCAATGCGCGGTACAGAGGTTCCCAACCCCGAGGGGGAGCTAATCTAGAAAGC	1257
Db	1201	TAAATCAATGCGCGGTACAGAGGTTCCCAACCCCGAGGGGGAGCTAATCTAGAAAGC	1260
QY	1258	GCGTCGTAAAGTCCGATCGAGAGTCGCAACTCGAATCCGTAAGTCGGATAGTAAT	1317
Db	1261	GCGTCGTAAAGTCCGATCGAGAGTCGCAACTCGAATCCGTAAGTCGGATAGTAAT	1320
QY	1318	CGCGGATCAGCAATGTCCGGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACAC	1377
Db	1321	CGCGGATCAGCAATGTCCGGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACAC	1380
QY	1378	CATGGGAATGGGGTTTCAACAGAAAGCAGTAGTCTAAACCTTAAGAGGGGCGCTTGCCACGG	1437
Db	1381	CATGGGAATGGGGTTTCAACAGAAAGCAGTAGTCTAAACCTTAAGAGGGGCGCTTGCCACGG	1440
QY	1438	TGAATTCATGACTGGGGTG	1457
Db	1441	TGAATTCATGACTGGGGTG	1460

[illegible]

```

Db      61 CTGGTGGCGAGTGGCGAAGCGGTGAGTAATGCATCGGAACGTATCCTTAAGTGGGGATA 120
Qy      121 AGCATCGAAGAAAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGCAGGGAGTCGAA 180
Db      121 AGCGATCGAAGAAAGATGGCGCTAATACCGCATATATCTCTGAGGAGAAAAGCAGGGAGTCGAA 180
Qy      181 GACCTTGCGCTTTTGGAGCGCGCGATGTCTGATTTAGCTAGTTGTTGGTGGGTTAAAGGCTTAC 240
Db      181 GACCTTGCGCTTTTGGAGCGCGCGATGTCTGATTTAGCTAGTTGTTGGTGGGTTAAAGGCTTAC 240
Qy      241 CAAAGGCGAGATCAATGATGTTGTTCTGAGAGAGACGACAGCAGACTGGGACTGAGACAG 300
Db      241 CAAAGGCGAGATCAATGATGTTGTTCTGAGAGAGACGACAGCAGACTGGGACTGAGACAG 300
Qy      301 GCCCAGACTCTCTACGGGAGGACAGAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360
Db      301 GCCCAGACTCTCTACGGGAGGACAGAGTGGGAAATTTTGGACAAATGGGCGGAAACCTGATC 360
Qy      361 CAGCAATGCCGCGTGAATGAAAGAGCGCTTTCGGGTGTAAAGCTCTTTCAGTGGAGAGA 420
Db      361 CAGCAATGCCGCGTGAATGAAAGAGCGCTTTCGGGTGTAAAGCTCTTTCAGCGGAAAGCA 420
Qy      421 AAAGCTTACGGTAAATATATGTCATCCATGACGGTATCGACAGAAAGAAAGACCGGCTTAC 480
Db      421 AAAGCTTACGGTAAATATATGTCATCCGATGACGGTATCGAGAAAGAAAGACCGGCTTAC 480
Qy      481 TAGCTGCCAGACGCGCGGTAAATACGTAGGGTGCACACGCTTAAATCGGAATTTACTGGGCGT 540
Db      481 TAGCTGCCAGACGCGCGGTAAATACGTAGGGTGCACACGCTTAAATCGGAATTTACTGGGCGT 540
Qy      541 AAAGGCTGCCAGCGCGCTTGTAAAGTCAATGTGAAATCCCGGGCTTAACTGGGAAT 600
Db      541 AAAGGCTGCCAGCGCGCTTGTGTAAAGTCAATGTGAAATCCCGGGCTTAACTGGGAAT 600
Qy      601 TGGCTTTGAAACTCAAAAGCTAAGTGTGGGAGAGGAGGTGAAATTCATGTGTACAG 660
Db      601 TGGCTTTGAAACTCAAAAGCTAAGTGTGGGAGAGGAGGTGAAATTCAGTGTACAG 660
Qy      661 TGAATTCGTAGATATGGAAGAACATTCATGTCGAAAGGACGCTCTGGGTTAACT 720
Db      661 TGAATTCGTAGATATGGAAGAACATTCATGTCGAAAGGACGCTCTGGGTTAACT 720
Qy      721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTTAGTATCTGTATGTCACGCG 780
Db      721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTTAGTATCTGTATGTCACGCG 780
Qy      781 CTAAACGATGCACTAGTTGTCCGGTCTTAAACGACGACTTGTAACTGATACGCGCTGA 840
Db      781 CTAAACGATGCACTAGTTGTCCGGTCTTAAACGACGACTTGTAACTGATACGCGCTGA 840
Qy      841 AGTTGACCCGCTGGGGAGTACGTCGCAAGATTAACCAAGGAAATGACGGGAGACC 900
Db      841 AGTTGACCCGCTGGGGAGTACGTCGCAAGATTAACCAAGGAAATGACGGGAGACC 900
Qy      901 GCACAAGCGGTGAGTTATGATGATTAATTCGATGCAACGCGCAAAACCTTACCTT 960
Db      901 GCACAAGCGGTGAGTTATGATGATTAATTCGATGCAACGCGCAAAACCTTACCTT 960
Qy      961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTGGGGAAGCTTAACACAGT 1017
Db      961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTGGGGAAGCTTAACACAGT 1017
Qy      1018 CTGCAATGCTGTCGTCAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1077
Db      1021 CTGCAATGCTGTCGTCAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1080
Qy      1078 ACCCTTGTCAATTAATTTGCGCATATTTGGTGGGCACTTTAAATGAGACTGCCGTGACAA 1137
Db      1081 ACCCTTGTCAATTAATTTGCGCATATTTCAATTGGGCACTTTAAATGAGAACTGCCGTGACAA 1140
Qy      1138 CCGAGAGAAAGTGGGAGTGAAGTCAAGTCTCATATGCGCTTATGGGTAGGGCTTCAACG 1197
Db      1141 CCGAGAGAAAGTGGGAGTGAAGTCAAGTCTCATATGCGCTTATGGGTAGGGCTTCAACG 1200

```

```

Qy      1198 TAATACATATGGCGCGTACAGAGGTTTCCCAACCCCGAGGGGAGCTAATCTCAAGAAC 1257
Db      1201 TAATACATATGGCGCGTACAGAGGTTTCCCAACCCCGAGGGGAGCTAATCTCAAGAAC 1260
Qy      1258 GCGTGTAGTCCGGAATCGAGTCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 1317
Db      1261 GCGTGTAGTCCGGAATCGAGTCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 1320
Qy      1318 CCGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGCTACAC 1377
Db      1321 CCGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGCTACAC 1380
Qy      1378 CATGGAGTGGGTTTACCAAGACAGTATGCTTAACCGTAAAGGAGGCGCTTGCCACG 1437
Db      1381 CATGGAGTGGGTTTACCAAGACAGTATGCTTAACCGGAAAGAGGCGCTTGCCACG 1440
Qy      1438 TGAGATTCACTGACTGGGCTG 1457
Db      1441 TGAGATTCACTGACTGGGCTG 1460

```

```

RESULT 39
AY123801 1497 bp DNA linear BCT 23-SEP-2003
LOCUS Nitrosospira sp. Nsp12 16S ribosomal RNA gene, partial sequence.
DEFINITION
AY123801
VERSION
AY123801.1 GI:24474430
KEYWORDS
SOURCE
ORGANISM

```

```

REFERENCE
1 (bases 1 to 1497)
Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roser,A. and
Koeps,H.P.
16S rRNA and amoa-based phylogeny of 12 novel betaproteobacterial
ammonia-oxidizing isolates: extension of the dataset and proposal
of a new lineage within the nitrosomonads
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1495-1494 (2003)

```

```

TITLE
JOURNAL
PUBMED
13130037
REFERENCE
2 (bases 1 to 1497)
Purkhold,U., Timmermann,G., Koeps,H.-P., Pommerening-Roeser,A. and
Wagner,M.
Direct Submission
Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany

```

```

FEATURES
source
1. .1497
/organism="Nitrosospira sp. Nsp12"
/mol_type="genomic DNA"
/strain="Nsp12"
/db_xref="taxon:136545"
<1..>1497
/product="16S ribosomal RNA"

```

```

ORIGIN
Query Match 89.3%; Score 1300.4; DB 1; Length 1497;
Best Local Similarity 94.1%; Pred. No. 2.2e-13;
Matches 1374; Conservative 0; Mismatches 81; Indels 5; Gaps 2;

```

```

Qy      1 ATTAAGCGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGACGGAATGCTTGAT 60
Db      3 ATTAAGCGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGACGGAATGCTTGAT 60
Qy      61 CTGGTGGCGATGGCGGACGGGTGATGATGATCGAGAAAGTATCCAGAAAGAGGGGGGTA 120
Db      61 CTGGTGGCGATGGCGGACGGGTGATGATGATGATGATGATGATGATGATGATGATGATG 120
Qy      121 ACAGATGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGATCGAA 180
Db      121 ACAGATGAAAGATGTGCTAATACCGCATATCTCTGAGAGAGAAAGCAGGGATCGCA 180

```

QY	181	AAACCTTGGCGCTTTTGGAGGGCGCCAGTGCCTGATTGAGCTAGTGTGGGTAAAGCCCTAC	240
Db	181	GACCTTGGCGCTTTTGGAGGGCGCCAGTGCCTGATTGAGCTAGTGTGGGTAAAGCCCTTAC	240
QY	241	CAAGCGACGAACTCAGTAGTGTGGTCTGAGAGAGCGACAGCCACACTGGGACTGAGACAG	300
Db	241	CAAGCGCTTCGATCAGTAGTGGTCTGAGAGAGCGACAGCCACACTGGGACTGAGACAG	300
QY	301	GCCCAAGACTCCTACCGGAGGCGACAGTGGGGAATTTGGACAATGGGCGCAAGCCTGATC	360
Db	301	GCCCAAGACTCCTACCGGAGGCGACAGTGGGGAATTTGGACAATGGGCGCAAGCCTGATC	360
QY	361	CAGCAATGCCGGGTAGAGAGAAGAAGGCTTCGGGTTGTAAACTCTTTCACTGCGAGAGA	420
Db	361	CAGCAATGCCGGGTAGAGAGAAGAAGGCTTCGGGTTGTAAACTCTTTCACTGCGAGAGA	420
QY	421	AAAGGTTACGGTAAATTAATCGTAGCCATGACGGTATCGACAGAGAAGACACCGGCTAAC	480
Db	421	AAAGGTTACGGTAAATTAATCGGTGACTAATGACGGTATCGGAAAGAAAGACACCGGCTAAC	480
QY	481	TACGTGCCAGCAGCCGCGGTAAATACGTAGAGGTGCAAACCGTTAATCGAATTACTGGGCGT	540
Db	481	TACGTGCCAGCAGCCGCGGTAAATACGTAGAGGTGCAGCGGTTAATCGAATTACTGGGCGT	540
QY	541	AAAGGCTGCGCAGCGCGCTTTGTAAGTCAGATGTGAATCCCGGCGTTAACTCTGGGAAT	600
Db	541	AAAGGCTGCGCAGCGCGGTTTGTAAATCGAATGTGAATCCCGGCGCTACAACTCTGGGAAC	600
QY	601	TGCGTTTGAAATATACAAAGCTAGAGTGTGGCGAGAGGAGTGGAAATTCATGTGTAGAG	660
Db	601	TGCGTTTGAAATTAACAAGCTAGAGTGTGGCGAGAGGAGTGGAAATTCACGTGTAGAG	660
QY	661	TGAAATGCGTAGAGATATGGAAGAACATCGATGCGGAAGCCGCTCTGGGTTAACT	720
Db	661	TGAAATGCGTAGAGATATGGAAGAACACCGATGCGGAAGCCGCTCTGGGTTAACT	720
QY	721	GACGCTCATGCAAGAAAGCGTGGGAGCAAAACGATTAAGTACCTGTGTAGTCCAGCC	780
Db	721	GACGCTCAGGCGACGAAAGCGTGGGAGCAAAACGATTAAGTACCTGTGTAGTCCAGCC	780
QY	781	CTAAACGATGTCAACTAGTGTGTGGGCTTAATTAAGGCTGTGTAAAGCTAAGCGGTGA	840
Db	781	CTAAACGATGTCAACTAGTGTGTGGGCTTTTAACGGACTTGTGTAAAGCTAAGCGGTGA	840
QY	841	AGTTGACCGGCTGCGGAGTAGTCCGTCGCAAGATTAAACCTCAAAAGGAATTGACGGGGAACC	900
Db	841	AGTTGACCGGCTGCGGAGTAGTCCGTCGCAAGATTAAACCTCAAAAGGAATTGACGGGGAACC	900
QY	901	GCAACAAGCGGTGATTAATGTGATTAATTCGATGCAAGCGGCAAAAACCTTAATCTACCTT	960
Db	901	GCAACAAGCGGTGATTAATGTGATTAATTCGATGCAAGCGGCAAAAACCTTAATCTACCTT	960
QY	961	GACATGTAGCGAATTTTCTAGAGATAGATTACTGC---TTGGGGAACGTTACAAGGTG	1011
Db	961	GACATGTAGCGAAGCCCGCCGAGAGGTAGTGTGCCGAAGAGGAACGGTTACACAGGTG	1022
QY	1018	CTGCATGACCTGTGCTCAGCTCGTGTGTAAGATGTGTGATTAAAGTCCGCAACGAGCCGA	1077
Db	1021	CTGCATGACCTGTGCTCAGCTCGTGTGTAAGATGTGTGATTAAAGTCCGCAACGAGCCGA	1088
QY	1078	ACCCTTGTCAATTAATTGCCATCATTTGTTGGGCACTTTAATGAGACTGCGCGGTACAAA	1137
Db	1081	ACCCTTGTCAATTAATTGCCATCATTTCAAGTGGGCACTTTAATGAACAATGCGCGGTACAAA	1144
QY	1138	CCGAGAGGAAGTGGGGATGACGTCAAGTCTCTATGGCCCTTATGGGTAGGCTTTACAGC	1197
Db	1141	CCGAGAGGAAGTGGGGATGACGTCAAGTCTCTATGAGCCCTTATGGGTAGGCTTTACACG	1200
QY	1198	TAAATCAATGGGCGCTTACAGAGGTTGCGCAACCCGCAAGGGGAGACTAATCTCAGAAAGC	1257
Db	1201	TAAATCAATGGGCGCTTACAGAGGTTGCGCAACCCGCAAGGGGAGACTAATCTCAGAAAGC	1266
QY	1258	GCGTGTAGTCCGATCGGAGTCTTCAACTCGATCTCGTGAAGTCGGAATCGTATGTAAT	1317

Db	1261	GGGTGTAGTCCGAGTCCGAGTCTGCAACTGCATCTCGTGAAGTGGATCGCTAGTAT	132
Qy	1318	CGCGGATCAGCATGTGCGGCGTGAATAGCTTCCGGGCTTTGTACACACCGCCGTCACAC	137
Db	1321	CGCGGATCAGCATGTGCGGCGTGAATAGCTTCCGGGCTTTGTACACACCGCCGTCACAC	1380
Qy	1378	CATGGAGTGGGTTTACACAGAAGGATAGCTTACCGTAAAGAGGGCGCTTGTGCACGG	143
Db	1391	CATGGAGTGGGTTTACACAGAAGGATAGCTTAAACCGCAAGAGGGGCGCTTGTGCACGG	144
Qy	1438	TGAGATTCATGACTCGGGGTG 1457	
Db	1441	TGAGATTCATGACTCGGGGTG 1460	
RESULT 40			
LOCUS	AY631270	1458 bp DNA linear BCT 28-DEC-200	
DEFINITION	Nitrosovibrio sp. FJ1423 16S ribosomal RNA gene, partial sequence.		
ACCESSION	AY631270		
VERSION	AY631270.1 GI:51093365		
KEYWORDS			
SOURCE	Nitrosovibrio sp. FJ1423		
ORGANISM	Nitrosovibrio sp. FJ1423		
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosovibrio.		
AUTHORS	1 (bases 1 to 1458)		
TITLE	Ida,T.; Satoh,M.; Yabe,R.; Takahashi,R. and Tokuyama,T.		
JOURNAL	Identification of Genus Nitrosovibrio, Ammonia-Oxidizing Bacteria, by Comparison of N-Terminal Amino Acid Sequences of Phosphoglycerate Kinase		
REFERENCE	J. Biosci. Bioeng. 98 (5), 380-383 (2004)		
AUTHORS	2 (bases 1 to 1458)		
TITLE	Takahashi,R., Yabe,R. and Tokuyama,T.		
JOURNAL	Nitrosovibrio sp. FJ1423 16S ribosomal RNA gene Unpublished		
REFERENCE	3 (bases 1 to 1458)		
AUTHORS	Takahashi,R., Yabe,R. and Tokuyama,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-MAY-2004) Agricultural and Biological Chemistry, College of Bioresource Sciences, Nihon University, 1866 Kameino, Fujiwara, Kanagawa 252-8510, Japan		
FEATURES	Location/Qualifiers		
source	1..1458		
	/organism="Nitrosovibrio sp. FJ1423"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:289540"		
	/clone="FJ1423"		
rRNA	<1..>1458		
	/product="16S ribosomal RNA"		
ORIGIN			
Query Match	89.0%; Score 1297.2; DB 1; Length 1458;		
Best Local Similarity	94.0%; Pred. No. 2.5e-13;		
Matches 1372; Conservative	0; Mismatches 83; Indels 5; Gaps 2;		
Qy	1	ATTGAACCTTGGCGGACGATCTTTACATATGCAATCGCAACGGCAGCAGGATGCTTGCAT	60
Db	1	ATTGAACCTTGGCGGACGATCTTTACATATGCAATCGCAACGGCAGCAGGATGCTTGCAT	58
Qy	61	CTGGTGGGAGTGGGCGGACGGGTAGTATCATCTCGAAGCTATTCAGAAAGGCGGAGTA	120
Db	59	CTGGTGGGAGTGGGCGGACGGGTAGTATCATCTCGAAGCTATTCCTTAAGTGGGGGATA	118
Qy	121	ACGCATCGAAAGATGTGCTAATATACCGCATATCTCTTAAGAGAGAAAGCAGGGAGTCGAA	180
Db	119	ACGCATCGAAAGATGTGCTAATATACCGCATATCTCTGAAGAGAAAGCAGGGAGTCGAA	178
Qy	181	GACCTTGGGCTTTTGGAGCGGCGCATGTCTGATTAGCTAAGTTGGTGGGTAAAGGCTTAC	240
Db	179	GACCTTGGGCTTTTGGAGCGGCGCATGTCTGATTAGCTAAGTTGGTGGGTAAAGGCTTAC	238

241 CAAGCGACGATCATAGTGTGCTGAGAGACGACCGACACTGGGACTGAGACAG 300  
239 CAAGGAGAGATCATAGTGTGCTGAGAGACGACCGACACTGGGACTGAGACAG 298  
301 GCCCAGACTCTTACGGGAGGACAGATGGGGAAATTTTGGACATATGGGCGCAACCTGATC 360  
299 GCCCAGACTCTTACGGGAGGACAGATGGGGAAATTTTGGACATATGGGCGCAACCTGATC 358  
361 CAGCAATGCCGCGTGAATGAAGAAGGCTTCGGGTTGTAAAGCTCTTTCAGTGAAGAGA 420  
359 CAGCCATGCCGCGTGAATGAAGAAGGCTTCGGGTTGTAAAGCTCTTTCAGCCGGAACGA 418  
421 AAAGTTACCGTAAATTAATCGTGAACCATGACGGTATCGACAGAAAGACACCGGCTAAC 480  
419 AAAGTTACCGTAAATTAATCGGTTGACTGACGTTACCGAAGAAAGACACCGGCTAAC 478  
481 TACGTGCACAGACCGCGGTAAATCGTAAAGGTTGCAAGCTTAAATTCGAGGCGCT 540  
479 TACGTGCACAGACCGCGGTAAATCGTAAAGGTTGCAAGCTTAAATTCGAGGCGCT 538  
541 AAAGGGTGGCAGGCGGCTTGTAAATCGATGTAATCCCGGCTTAACTGGGAAT 600  
539 AAAGCGGCGCAGGCGGCTTGTAAATCGATGTAATCCCGGCTTAACTGGGAAT 598  
601 TGGCTTTGAACCTAACAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTACAG 660  
599 TGGCTTTGAACCTAACAGCTAGAGTGTGACAGAGGAGGAGGTGGAATTCATGTGTACAG 658  
661 TGAATTCGTAGATGTGAAGAAACATCGATGCGAAGGACGCTCTCGGTTAACTACT 720  
659 TGAATTCGTAGATGTGAAGAAACATCGATGCGAAGGACGCTCTCGGTTAACTACT 718  
721 GAGCTCATCAAGAAAGCGTGGGAGCAACAGATTAAGTACCTGTGTATCCAGGCC 780  
719 GAGCTCATCAAGAAAGCGTGGGAGCAACAGATTAAGTACCTGTGTATCCAGGCC 778  
781 CTAAACGATGCACTAGTGTGTGGGCTTATTAAGGCTTGTGTAAAGAGCTAACGCGTGA 840  
779 CTAAACGATGCACTAGTGTGTGGGCTTATTAAGGCTTGTGTAAAGAGCTAACGCGTGA 838  
841 AGTTGACCGCTGGGAGAGTACGTCGCAAGATTTAAACTCAAAGAAATTAACCGGGAGCC 900  
839 AGTTGACCGCTGGGAGAGTACGTCGCAAGATTTAAACTCAAAGAAATTAACCGGGAGCC 898  
901 GCAAGCGGTGATTAATGTGATTAATGTGATGAACGCGAAGAACTTACTACTACCTT 960  
899 GCAAGCGGTGATTAATGTGATTAATGTGATGAACGCGAAGAACTTACTACTACCTT 958  
961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC--TTGCGGAACGCTTAACAAGGTG 1017  
959 GACATGTAGCGAACCGCGCTAGAGAGTGTGTCGCCAAGAGGAGCGGTAAACAAGGTG 1018  
1018 CTGATGAGCTGTCTGACGTCTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1077  
1019 CTGATGAGCTGTCTGACGTCTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1078  
1078 ACCCTTGTCAATTAATGGCATCTTGGTGGGCACTTAATGAAGCTGCGGTTGACAA 1137  
1079 ACCCTTGTCAATTAATGGCATCTTGGTGGGCACTTAATGAAGCTGCGGTTGACAA 1138  
1138 CCGAGAGAGGTGGGAGTGAAGTCTCAATGACCCTTAATGGGTTGACAA 1197  
1139 CCGAGAGAGGTGGGAGTGAAGTCTCAATGACCCTTAATGGGTTGACAA 1198  
1199 TAAATCAATGGCGCTTACAGAGGGTTGCCAACCCGAGAGGGGAGCTAATCTCAGAAAAC 1258  
1258 GCGTGTAGTCCGGATTCGGAATCTGCACTCCGTTGAAGTGGGAATGCTAGTAAAT 1317  
1259 GCGTGTAGTCCGGATTCGGAATCTGCACTCCGTTGAAGTGGGAATGCTAGTAAAT 1318  
1318 CCGGATCAGCATGTCCGGTGAATAGTTCGCCGGGCTTGTACAACCGCCCTCAGAC 1377

1319 CCGGATCAGCATGTCCGGTGAATAGTTCGCCGGTCTTGTACACACCGCCGCTCAGAC 1378  
1378 CATGGAGTGGGTTTTCACAGAGAGAGTGTCTAAACGTTAAAGAGAGGCGCTTGGCAGG 1437  
1379 CATGGAGTGGGTTTTCACAGAGAGAGTGTGTAAACCGCAAGAGAGGCGCTTGGCAGG 1438  
1438 TGAATTCATGACTGGGCTG 1457  
1439 TGAATTCATGACTGGGCTG 1458  
RESULT 41  
AY631269 1458 bp DNA linear BCT 28-DEC-2004  
LOCUS Nitrosovibrio sp. FJ182 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY631269  
ACCESSION AY631269  
VERSION AY631269.1 GI:51093364  
KEYWORDS  
SOURCE  
ORGANISM Nitrosovibrio sp. FJ182  
Nitrosovibrio sp. FJ182  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosovibrio.  
REFERENCE  
AUTHORS Ida,T., Satoh,M., Yabe,R., Takahashi,R. and Tokuyama,T.  
TITLE Identification of Genus Nitrosovibrio, Ammonia-oxidizing Bacteria,  
by Comparison of N-Terminal Amino Acid Sequences of  
Phosphoglycerate Kinase  
J. Biosci. Bioeng. 98 (5), 380-383 (2004)  
REFERENCE  
AUTHORS Takahashi,R., Sato,N. and Tokuyama,T.  
TITLE Nitrosovibrio sp. FJ182 16S ribosomal RNA gene  
JOURNAL Unpublished  
AUTHORS 3 (bases 1 to 1458)  
Takahashi,R., Sato,N. and Tokuyama,T.  
TITLE Direct Submission  
JOURNAL Submitted (20-MAY-2004) Agricultural and Biological Chemistry,  
College of Bioresource Sciences, Nihon University, 1866 Kameino,  
Fujisawa, Kanagawa 252-8510, Japan  
LOCATION/Qualifiers  
FEATURES  
source  
1..1458  
/organism="Nitrosovibrio sp. FJ182"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:289539"  
/clone="FJ182"  
<1..>1458  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 88.7%; Score 1292.4; DB 1; Length 1458;  
Best Local Similarity 93.8%; Pred. No. 3e-13;  
Matches 1369; Conservative 0; Mismatches 86; Indels 5; Gaps 2;  
1 ATTAAGCGTGGCGGATGCTTTTACATCATCAAGTGAAGGAGCAGCGATGCTTCAT 60  
2 ATTGAAGCGTGGCGGATGCTTTTACATCATCAAGTGAAGGAGCAGCGATGCTTCAT 58  
1 ATTGAAGCGTGGCGGATGCTTTTACATCATCAAGTGAAGGAGCAGCGATGCTTCAT 58  
59 CTGTGCGGATGGCGGAACGGGTGAATGATGATCGAAGCTATCTTAAGTGGGGATA 118  
61 CTGTGCGGATGGCGGACGGGTGAATGATGATCGAAGCTATCTTAAGTGGGGATA 120  
CTGTGCGGATGGCGGACGGGTGAATGATGATCGAAGCTATCTTAAGTGGGGATA 120  
59 CTGTGCGGATGGCGGAACGGGTGAATGATGATCGAAGCTATCTTAAGTGGGGATA 118  
121 ACGCATGAAGATGTCTAATACCGCATATCTTAAAGAGAAAGCAGGGATCGAAA 180  
119 ACGCATGAAGATGTCTAATACCGCATATCTTAAAGAGAAAGCAGGGATCGAAA 178  
181 GACCTTGGCTTTTGGAGCGGCGCATGCTGATTAAGTGAAGTGGGTTAAAGCCCTAC 240  
179 GACCTTGGCTTTTGGAGCGGCGCATGCTGATTAAGTGAAGTGGGTTAAAGCCCTAC 238  
241 CAAGGCGCATCATGATGTTGTTGTGAAGAGCAGCAGCACACTGGGACTGAGACAG 300  
239 CAAGGCGCATCATGATGTTGTTGTGAAGAGCAGCAGCACACTGGGACTGAGACAG 298

301 GCCCAGACTCTCTACGGGAGGACAGAGTGGGAAATTTTGGACAAATGGGCGCAGCCTGATC 360  
DB 299 GCCCAGACTCTCTACGGGAGGACAGAGTGGGAAATTTTGGACAAATGGGCGCAGCCTGATC 358  
QY 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCCGGGTTGTAAAGCTCTTTCAGTGAAGAGA 420  
DB 359 CAGCAGTGGCGGTGAAGTGAAGAGGCTTCCGGGTTGTAAAGCTCTTTCAGTGAAGAGA 418  
QY 421 AAGGTTACGGTAAATATATGTGACCCATGACGGTATCCAGAGAAAGACCGGCTAAC 480  
DB 419 AAGGTTACGGTAAATATATGTGACCCATGACGGTATCCAGAGAAAGACCGGCTAAC 478  
QY 481 TAGTGCACGACCCCGGTAAATACGATGGGTGAAGCGTTAATCCGAATTAATCTGGGCGT 540  
DB 479 TAGTGCACGACCCCGGTAAATACGATGGGTGAAGCGTTAATCCGAATTAATCTGGGCGT 538  
QY 541 AAGGTTGCGCAGGCGGCTTGTAAAGTCAAGATGTGAATCCCGGCTTAACCTGGGAAT 600  
DB 539 AAGGTTGCGCAGGCGGCTTGTAAAGTCAAGATGTGAATCCCGGCTTAACCTGGGAAT 598  
QY 601 TGGCTTTGAAATCTACAAAGCTTAAGTGTGGCAGAGGAGGTGAAATTCATGTGTAGCAG 660  
DB 599 TGGCTTTGAAATCTACAAAGCTTAAGTGTGGCAGAGGAGGTGAAATTCATGTGTAGCAG 658  
QY 661 TGAATATGCGTGAAGATATGGAAGACATCGATGCGCAAGGAGCAGCTCTGGGTTAACT 720  
DB 659 TGAATATGCGTGAAGATATGGAAGACATCGATGCGCAAGGAGCAGCTCTGGGTTAACT 718  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGTATGTCACGCC 780  
DB 719 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGTATGTCACGCC 778  
QY 781 CTAAACGATGTCACTAATGTTGTGGGCTTATTAAGCTTGTGTACGAAGCTTAACGGGTGA 840  
DB 779 CTAAACGATGTCACTAATGTTGTGGGCTTATTAAGCTTGTGTACGAAGCTTAAACGGGTGA 838  
QY 841 AGTTGACCGCTGGGAGTACGGTTCGCAAGATTTAAACTCAAGAAATTGACCGGAGACC 900  
DB 839 AGTTGACCGCTGGGAGTACGGTTCGCAAGATTTAAACTCAAGAAATTGACCGGAGACC 898  
QY 901 GCACAGCGGCTGATTAATGTGATTAATTCGATGCAACCGGAAACCTTACCTACCTT 960  
DB 899 GCACAGCGGCTGATTAATGTGATTAATTCGATGCAACCGGAAACCTTACCTACCTT 958  
QY 961 GACATGTAGCAATTTTCTAGAGATAGATTAGTC---TTCCGGAACGCTTAACACAGGTC 1017  
DB 959 GACATGTAGCAATTTTCTAGAGATAGATTAGTC---TTCCGGAACGCTTAACACAGGTC 1018  
QY 1018 CTGCATGGGCTGTGTAGGCTGTGTGTGATGTTGGGTTAGTCCCGCAACGAGCGCA 1077  
DB 1019 CTGCATGGGCTGTGTAGGCTGTGTGTGATGTTGGGTTAGTCCCGCAACGAGCGCA 1078  
QY 1078 ACCCTTGTCAATTAATTTGATTCATTTTGGTGGCAGCTTTAATGAGACTGCGGTGACAA 1137  
DB 1079 ACCCTTGTCAATTAATTTGATTCATTTTGGTGGCAGCTTTAATGAGACTGCGGTGACAA 1138  
QY 1138 CCGGAGAAAGGTGGGAGTGAAGTCAAGTCTCATGCGCCCTTAATGGGTAGGGCTTCAACG 1197  
DB 1139 CCGGAGAAAGGTGGGAGTGAAGTCAAGTCTCATGCGCCCTTAATGGGTAGGGCTTCAACG 1198  
QY 1198 TAAATCAATGGCGCGTACAGAGGGTTGGCAACCGGAGGGGAGGCTAATCTCAGAAAGC 1257  
DB 1199 TAAATCAATGGCGCGTACAGAGGGTTGGCAACCGGAGGGGAGGCTAATCTCAGAAAGC 1258  
QY 1258 GCGTCTAGTCCGAGTCCGAGTCTGCAACTCGACTCGTGAATCGGAATCGCTAGTAAT 1317  
DB 1259 GCGTCTAGTCCGAGTCCGAGTCTGCAACTCGACTCGTGAATCGGAATCGCTAGTAAT 1318  
QY 1318 CCGGAGTCAAGATGTGCGGCTGAATACGTTCCGGGCTTTGTACACACCGCCCGTCAAC 1377  
DB 1319 CCGGAGTCAAGATGTGCGGCTGAATACGTTCCGGGCTTTGTGTACACACCGCCCGTCAAC 1378  
QY 1378 CATGGAGTGGGTTTACCAAGAAAGCATAGTCTAAACCGTAAAGAGGGCGCTTGCCACGG 1437

DB 1379 CATGGAGTGGGTTTACCAAGAAAGCATAGTCTAAACCGTAAAGAGGGCGCTTGCCACGG 1438  
QY 1438 TGAATTAATGACTGGGCTG 1457  
DB 1439 TGAATTAATGACTGGGCTG 1458  
RESULT 42  
AY123808 1497 bp DNA linear BCT 23-SEP-2003  
LOCUS Nitrosospira sp. Nsp1 16S ribosomal RNA gene, partial sequence.  
DEFINITION  
ACCESSION AY123808  
VERSION AY123808.1 GI:24474437  
KEYWORDS  
SOURCE Nitrosospira sp. Nsp1  
ORGANISM Nitrosospira sp. Nsp1  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira.  
REFERENCE  
1 Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roesser, A. and  
Koops, H.-P. 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
JOURNAL  
PUBMED 13130037  
REFERENCE  
2 (bases 1 to 1497)  
Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roesser, A. and  
Wagner, M.  
TITLE  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1. 1497  
/organism="Nitrosospira sp. Nsp1"  
/mol\_type="genomic DNA"  
/strain="Nsp1"  
/db\_xref="taxon:136547"  
<1. >1497  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 88.7%; Score 1292.4; DB 1; Length 1497;  
Best Local Similarity 93.8%; Pred. No. 2.9e-13;  
Matches 1369; Conservative 0; Mismatches 86; Indels 5; Gaps 2;  
QY 1 ATTGAACGCTGGCGGATGCTTTTACATGCAAGTCAACGCGACAGCGATGCTTGCAT 60  
DB 3 ATTGAACGCTGGCGGATGCTTTTACATGCAAGTCAACGCGACAGCGATGCTTGCAT 60  
QY 61 CTGTGGCGGAGTGGCGGAGCGGTGAGTAATGATGGAACGTAATCAAGAAAGGGGCTA 120  
DB 61 CTGTGGCGGAGTGGCGGAGCGGTGAGTAATGATGGAACGTAATCAAGAAAGGGGCTA 120  
QY 121 AGCAGTCCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGAGATCGAA 180  
DB 121 AGCAGTCCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGAGATCGAA 180  
QY 181 GACCTTGGCTTTTGGAGCGCGCATGCTGATTAATGCTAGTGTGGGTAAAGCGCTAC 240  
DB 181 GACCTTGGCTTTTGGAGCGCGCATGCTGATTAATGCTAGTGTGGGTAAAGCGCTAC 240  
QY 241 CAAAGGAGATCAAGTGTGTCTGAGAGACGACACGACCACTGGGACTGAGACAGC 300  
DB 241 CAAAGGAGTCAAGTGTGTCTGAGAGACGACACGACCACTGGGACTGAGACAGC 300  
QY 301 GCCGAGACTCTTACGGGAGGCGAGTGGGAAATTTTGGACAAATGGGCGCAAGCCTGATC 360  
DB 301 GCCGAGACTCTTACGGGAGGCGAGTGGGAAATTTTGGACAAATGGGCGCAAGCCTGATC 360  
QY 361 CAGCAATGCGGCTGATGTAAGAGGCTTCCGGTGTAAAGCTTTTCAGTGAAGAGA 420

Db 361 CAGCCATGCGCGTGAATGAAGAAGCCCTTCGGGTTGTAAAGCTCTTTACGCCGAAACGA 420  
Qy 421 AAGGTTACGGTAAATATATGATACCATATGACGTTATCGACAGAAAGACACCGGCTAAC 480  
Db 421 AAGGTTACGGTAAATATGATACCATATGACGTTATCGACAGAAAGACACCGGCTAAC 480  
Qy 481 TACGTCACAGACACCGGCTAAATACGTAAGGTCGCAACCGTTAATCGGAATTTACTGGCGCT 540  
Db 481 TACGTCACAGACACCGGCTAAATACGTAAGGTCGCAACCGTTAATCGGAATTTACTGGCGCT 540  
Qy 541 AAGGTCGCGACGCGGCTTTGTAAAGTCAGATGTGAATCCCGGCTCAACCTGGAGAC 600  
Db 541 AAGGTCGCGACGCGGCTTTGTAAAGTCAGATGTGAATCCCGGCTCAACCTGGAGAC 600  
Qy 601 TGGGTTTGAACCTACAAAGCTAGAGTGTGCGACAGGAGGTGGAATTCATGTGTAGCAG 660  
Db 601 TGGGTTTGAACCTACAAAGCTAGAGTGTGCGACAGGAGGTGGAATTCATGTGTAGCAG 660  
Qy 661 TGAATATGCTAGATATGGAAGAACATCGATGCGAAGGCAAGCTTCGCGTTAACACT 720  
Db 661 TGAATATGCTAGATATGGAAGAACATCGATGCGAAGGCAAGCTTCGCGTTAACACT 720  
Qy 721 GAGGCTCATGACGAAGACGCTGGGAGCAACAGGATTTAGATACCTCTGTATGTCACGCC 780  
Db 721 GAGGCTCATGACGAAGACGCTGGGAGCAACAGGATTTAGATACCTCTGTATGTCACGCC 780  
Qy 781 CTAAAGCATGTCATCTAGTTGTTGGGCTTATAGGCTTGGTAAAGAACTAACGCGCTGA 840  
Db 781 CTAAAGCATGTCATCTAGTTGTTGGGCTTATAGGCTTGGTAAAGAACTAACGCGCTGA 840  
Qy 841 AGTTGACCGCTGGGAGATACGCTCGCAAGATTAAACTCAAGAAATTGACGGGAGACC 900  
Db 841 AGTTGACCGCTGGGAGATACGCTCGCAAGATTAAACTCAAGAAATTGACGGGAGACC 900  
Qy 901 GCACAGCGGTGATTTATGTGATTTAATTCATGCAAGCGCAAAACCTTACCTACCTT 960  
Db 901 GCACAGCGGTGATTTATGTGATTTAATTCATGCAAGCGCAAAACCTTACCTACCTT 960  
Qy 961 GACATGTAGGGAATTTCTAGAGATAGATTAGTGC---TTGGGAGCGCTAACAGGTCG 1017  
Db 961 GACATGTAGGGAATTTCTAGAGATAGATTAGTGC---TTGGGAGCGCTAACAGGTCG 1017  
Qy 1018 CTGATGAGGCTGTCTGACGCTGCTGTCGTGATGATGTGGGTTAAGTCCCGCAAGACGCA 1077  
Db 1018 CTGATGAGGCTGTCTGACGCTGCTGTCGTGATGATGTGGGTTAAGTCCCGCAAGACGCA 1077  
Qy 1021 CTGATGAGGCTGTCTGACGCTGCTGTCGTGATGATGTGGGTTAAGTCCCGCAAGACGCA 1080  
Db 1021 CTGATGAGGCTGTCTGACGCTGCTGTCGTGATGATGTGGGTTAAGTCCCGCAAGACGCA 1080  
Qy 1078 ACCCTTGTCATTAATTGTCATCATTTGTTGGGCACTTTAATGAGACTGCGGTCGACAA 1137  
Db 1078 ACCCTTGTCATTAATTGTCATCATTTGTTGGGCACTTTAATGAGACTGCGGTCGACAA 1137  
Qy 1081 ACCCTTGTCATTAATTGTCATCATTTGTTGGGCACTTTAATGAGACTGCGGTCGACAA 1140  
Db 1081 ACCCTTGTCATTAATTGTCATCATTTGTTGGGCACTTTAATGAGACTGCGGTCGACAA 1140  
Qy 1138 CCGAGAGAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1197  
Db 1138 CCGAGAGAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1197  
Qy 1141 CCGAGAGAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
Db 1141 CCGAGAGAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
Qy 1198 TAATACATGCGCGGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1257  
Db 1198 TAATACATGCGCGGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1257  
Qy 1201 TAATACATGCGCGGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1260  
Db 1201 TAATACATGCGCGGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1260  
Qy 1258 GCGTCGTAGTCCGAGTCGAGAGCTGCAACTGCACTCCGTGAAGTGGGAATGCTAGTAT 1317  
Db 1258 GCGTCGTAGTCCGAGTCGAGAGCTGCAACTGCACTCCGTGAAGTGGGAATGCTAGTAT 1317  
Qy 1261 GCGTCGTAGTCCGAGTCGAGAGCTGCAACTGCACTCCGTGAAGTGGGAATGCTAGTAT 1320  
Db 1261 GCGTCGTAGTCCGAGTCGAGAGCTGCAACTGCACTCCGTGAAGTGGGAATGCTAGTAT 1320  
Qy 1318 CCGGATCAGCATGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377  
Db 1318 CCGGATCAGCATGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377  
Qy 1321 CCGGATCAGCATGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
Db 1321 CCGGATCAGCATGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
Qy 1378 CATGGAATGAGGTTTACCAAGAGAGGATGTCTAACCGTAAAGAGAGGCGCTTGCACGG 1437  
Db 1378 CATGGAATGAGGTTTACCAAGAGAGGATGTCTAACCGTAAAGAGAGGCGCTTGCACGG 1437  
Qy 1381 CATGGAATGAGGTTTACCAAGAGAGGATGTCTAACCGTAAAGAGAGGCGCTTGCACGG 1440  
Db 1381 CATGGAATGAGGTTTACCAAGAGAGGATGTCTAACCGTAAAGAGAGGCGCTTGCACGG 1440  
Qy 1438 TGAATTCATGACTGGGTCG 1457  
Db 1438 TGAATTCATGACTGGGTCG 1457  
Qy 1441 TGAATTCATGACTGGGTCG 1460  
Db 1441 TGAATTCATGACTGGGTCG 1460

RESULT 43  
AY123813 1497 bp DNA linear BCT 23-SEP-2003  
LOCUS Nitrosospira sp. Nsp65 16S ribosomal RNA gene, partial sequence.  
DEFINITION  
ACCESSION AY123813  
VERSION AY123813.1 GI:24474442  
KEYWORDS  
SOURCE Nitrosospira sp. Nsp65  
ORGANISM Nitrosospira sp. Nsp65  
REFERENCE  
AUTHORS Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Rosser,A. and  
Koops,H.P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrospinae  
JOURNAL Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
PUBMED 13130037  
REFERENCE 2 (bases 1 to 1497)  
AUTHORS Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Rosser,A. and  
Wagner,M.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hoechanger 4, Freising 85350, Germany  
FEATURES  
source  
1..1497  
/organism="Nitrosospira sp. Nsp65"  
/mol\_type="genomic DNA"  
/strain="Nsp65"  
/db\_xref="taxon:200121"  
<1..>1497  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 88.7%; Score 1292.4; DB 1; Length 1497;  
Best Local Similarity 93.8%; Pred. No. 2; 9e-13;  
Matches 1369; Conservative 0; Mismatches 86; Indels 5; Gaps 2;  
Qy 1 ATTGAAGCGTGGGGGATGCTTTACATGATCAAGTCGACGCGACGATGCTTCAT 60  
Db 3 ATTGAAGCGTGGGGGATGCTTTACATGATCAAGTCGACGCGACGATGCTTCAT 60  
Qy 61 CTGGTGGCGAGTGGCGGACGCGTGAATGATGATCGGAACGTATCCAGAAAGGGGCTA 120  
Db 61 CTGGTGGCGAGTGGCGGACGCGTGAATGATGATCGGAACGTATCCAGAAAGGGGCTA 120  
Qy 121 ACCGATGAAGATGTGCTTAATCCGATATCTCTTAAGAGAAAGCAGGGATTCGAA 180  
Db 121 ACCGATGAAGATGTGCTTAATCCGATATCTCTTAAGAGAAAGCAGGGATTCGAA 180  
Qy 181 GACCTTCGCTTTTGGAGCGGCGGATGCTGATTAAGTGTGGGGTAAAGGCTTAC 240  
Db 181 GACCTTCGCTTTTGGAGCGGCGGATGCTGATTAAGTGTGGGGTAAAGGCTTAC 240  
Qy 241 CAAAGGACGATCAGTATGTTGCTGAGAGGACGACGACCACTGGGACTGAGACAG 300  
Db 241 CAAAGGACGATCAGTATGTTGCTGAGAGGACGACGACCACTGGGACTGAGACAG 300  
Qy 301 GCCCAACTCTTACGGAAGGACGACGATGAGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360  
Db 301 GCCCAACTCTTACGGAAGGACGACGATGAGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360  
Qy 361 CAGCAATGCGCGGAGTGAAGAGGCTTGGGTTGTAAGCTCTTTCACTGAGAGA 420  
Db 361 CAGCAATGCGCGGAGTGAAGAGGCTTGGGTTGTAAGCTCTTTCACTGAGAGA 420  
Qy 421 AAGGTTACGGTAAATATGCTGAACCATGACGCTATGACAGAAAGACACCGGCTAAC 480  
Db 421 AAGGTTACGGTAAATATGCTGAACCATGACGCTATGACAGAAAGACACCGGCTAAC 480



481 TACGTCCAGACCGCGGTAATACGTAGGGGTGCAAGCGTTAATCGGAATTACTGGGCGT 540  
481 TACGTCCAGACCGCGGTAATACGTAGGGGTGCAAGCGTTAATCGGAATTACTGGGCGT 540  
541 AAAAGGTGCGCAGCGCGCTTGTAACTCAGATGTGAATCCCGGGCTTAACTGGGAAT 600  
541 AAAAGGTGCGCAGCGCGCTTGTAACTCAGATGTGAATCCCGGGCTTAACTGGGAAT 600  
601 TGGGTTTGAACCTCAACAAAGCTAGATGTGTGCAAGGAGGTGGAATTCATGTGTGACG 660  
601 TGGGTTTGAACCTCAACAAAGCTAGATGTGTGCAAGGAGGTGGAATTCATGTGTGACG 660  
661 TGAATATGCGTAGATATGAGAAATCATCATGATGCGCAAGGACGCTCTGGGTTAACT 720  
661 TGAATATGCGTAGATATGAGAAATCATCATGATGCGCAAGGACGCTCTGGGTTAACT 720  
721 GACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTAGATACCTGTGTGACGCGC 780  
721 GACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTAGATACCTGTGTGACGCGC 780  
781 CTAAAGATGTCACTAGTTGTGGGCTTATTAAGCTTGTATGCAAGAGCTAACGGGTGA 840  
781 CTAAAGATGTCACTAGTTGTGGGCTTATTAAGCTTGTATGCAAGAGCTAACGGGTGA 840  
841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAACTCAAGAAATTGACGGGAGCC 900  
841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAACTCAAGAAATTGACGGGAGCC 900  
901 GCAAGAGCGGTGATTAATGTGATTAATTCGATGCAACGCAAAACCTTACCTACCTT 960  
901 GCAAGAGCGGTGATTAATGTGATTAATTCGATGCAACGCAAAACCTTACCTACCTT 960  
961 GACATGTACGGAATTTCTAGAGATAGATTAGTC---TTCCGGGAACGCTAACACAGG 1017  
961 GACATGTACGGAATTTCTAGAGATAGATTAGTC---TTCCGGGAACGCTAACACAGG 1017  
1018 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
1018 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
1021 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1021 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1078 ACCCTTGTCAATTAATTCGATTAATTCGATTAATTCGATTAATTCGATTAATTCGATTA 1137  
1081 TCCCTTGTCAATTAATTCGATTAATTCGATTAATTCGATTAATTCGATTAATTCGATTA 1140  
1138 CCGGAGGAAGTGGGAGTACGCTGCAAGCTGCTGCAAGCTGCTGCAAGCTGCTGCAAGCT 1197  
1138 CCGGAGGAAGTGGGAGTACGCTGCAAGCTGCTGCAAGCTGCTGCAAGCTGCTGCAAGCT 1197  
1141 CCGGAGGAAGTGGGAGTACGCTGCAAGCTGCTGCAAGCTGCTGCAAGCTGCTGCAAGCT 1200  
1141 CCGGAGGAAGTGGGAGTACGCTGCAAGCTGCTGCAAGCTGCTGCAAGCTGCTGCAAGCT 1200  
1198 TAATCAATGCGCTGCAAGAGGCTTGCACACCGGAGGAGGAGCTTAATTCAGAAAC 1257  
1201 TAATCAATGCGCTGCAAGAGGCTTGCACACCGGAGGAGGAGCTTAATTCAGAAAC 1260  
1258 GCGTGTAGTCCGATCGAGATCGCACTGCTGCAAGCTGCTGCAAGCTGCTGCAAGCT 1317  
1261 GCGTGTAGTCCGATCGAGATCGCACTGCTGCAAGCTGCTGCAAGCTGCTGCAAGCT 1320  
1318 CCGGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1377  
1321 CCGGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
1378 CATGGAGTGGGTTTCAACCAAGAGCAGTATCTAACCGTAAAGAGGAGGCTTGCACAG 1437  
1381 CATGGAGTGGGTTTCAACCAAGAGCAGTATCTAACCGTAAAGAGGAGGCTTGCACAG 1440  
1438 TGAGATTCTGACTGGGCTG 1457  
1441 TGAGATTCTGACTGGGCTG 1460

RESULT 44  
AJ621028 1462 bp DNA linear BCT 08-JUN-2004  
LOCUS AJ621028  
DEFINITION Nitrosomonas sp. 1s86 16S rRNA gene, isolate 1s86.

ACCESSION AJ621028  
VERSION AJ621028.1 GI:40994847  
KEYWORDS 16S ribosomal RNA; 16S rRNA gene.  
SOURCE Nitrosomonas sp. 1s86  
ORGANISM Nitrosomonas sp. 1s86  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosomonas.  
REFERENCE  
1 Nielsen, S., Beer Gillissen, M.J., Revsbech, N., Laanbroek, H.J. and Bollmann, A. Influence of heterotrophic bacteria on the activity and growth behavior of ammonia-oxidizing bacteria Unpublished  
2 (bases 1 to 1462)  
Bollmann, A.  
Direct Submission  
Submitted (15-JUN-2004) Bollmann A., Department for Microbial Ecology, NIOO-KNAW Centre for Limnology, Rijksstratweg 6, 3631 AC Nieuwersluis, NETHERLANDS  
FEATURES  
source  
1..1462  
/organism="Nitrosomonas sp. 1s86"  
/mol\_type="genomic DNA"  
/isolate="1s86"  
/isolation\_source="fresh water sediment"  
/db\_xref="taxon:261294"  
/country="Netherlands;Lake Drontmeer"  
1..1462  
/gene="16S rRNA"  
1..1462  
/gene="16S rRNA"  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 88.7%; Score 1292; DB 1; Length 1462;  
Best Local Similarity 94.4%; Pred. No. 3; le-13;  
Matches 1372; Conservative 0; Mismatches 76; Indels 5; Gaps 3;  
1 ATTTGAACCTGCGGCGCATGCTTTACATGCAAGTCGAACCGGACAGGATGCTTGAT 60  
12 ATTGAACCTGCGGCGCATGCTTTACATGCAAGTCGAACCGGACAGGATGCTTGAT 71  
61 CTGTGGGAGTGGCGGACGCGGTGAGTATGATCGAAGCGTATCAGAAAGGCGGGTA 120  
72 CTGTGGGAGTGGCGGACGCGGTGAGTATGATCGAAGCGTATCAGAAAGGCGGGTA 131  
121 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGCAAGGAGTCGAAA 180  
132 ACGCATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGTGGGGATCCAAA 191  
181 GACCTTGGCTTTTGGACCGCGCATGCTGATTAAGTGTGTTGGGGTAAAGGCTTAC 240  
192 AGCTTCATGCTTTTGGACCGCGCATGCTGATTAAGTGTGTTGGGGTAAAGGCTTAC 251  
241 CAAGCGACGATCAGTATGCTGAGAGGAGCAACGACACACTGGGACTGAGACAG 300  
252 CAAGCGACGATCAGTATGCTGAGAGGAGCAACGACACACTGGGACTGAGACAG 311  
301 GCCCACTCTTCAACGAGGAGCAGCAGTGGGAAATTTTGAACATGAGCGCAAGCTGATC 360  
312 GCCCACTTCTTCAACGAGGAGCAGCAGTGGGAAATTTTGAACATGAGCGCAAGCTGATC 371  
361 CAGCAATGCGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
372 CAGCAATGCGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 430  
421 AAAAGTTACGGTAAATATCGTACCCATGACGATGACGATGACGAGAAAGACACCGGCTAAC 480  
431 AAAAATTTGGCTAATATCCAGAAATTCATGACGATGACGAGAAAGACACCGGCTAAC 490  
481 TACGTCCAGACGCGCGGTATATACGTAGGGTGAAGGCTTAAATCGAATTAATCTGGGCGT 540  
491 TACGTCCAGACGCGCGGTATATACGTAGGGTGAAGGCTTAAATCGAATTAATCTGGGCGT 550



```

QY 541 AAAGGTGCGACGCGCCCTTGTAGTCAGATGTGAATCCCGGGCTTAACCTGGAAAT 600
Db 551 AAAGGTGCGACGCGCTTGTGTATGATGATGAAATCCCGGGCTTAACCTGGAAAT 610
QY 601 TGGGTTTGAATCAAAAGCTAGATGTGCGACAGGGAGGTGGAATTTCCATGTGTACAG 660
Db 611 TGGGTTTGAATCAAAAGCTAGATGTGCGACAGGGAGGTGGAATTTCCATGTGTACAG 670
QY 661 TGAATATCCGTAGATGTGAAGAACAATGATGCGAGGACCTCCCTGGTTAACT 720
Db 671 TGAATATCCGTAGATGTGAAGAACAATGATGCGAGGACCTCCCTGGTTAACT 730
QY 721 GACGCTCATGACGAAAGCTGGGGAGCAAAAGATTAGATCCTGTAGTCCAGCC 780
Db 731 GACACTCAGCAGAAAGCTGGGGAGCAAAAGATTAGATCCTGTAGTCCAGCC 790
QY 781 CTAAACGATGCACTAGTTGTTGGCCCTTATTAGGCTTGTATCCAGAGCTTAACGGTGA 840
Db 791 CTAACTATGTCAACTAGTTGTTGGCTTATTAGACTTGTATCCAGAGCTTAACGGTGA 850
QY 841 AGTTGACCCGCTGGGGAGTACGGTCCGAAGATTAAACTCAAGGAATTGACGGGGACCC 900
Db 851 AGTTGACCCGCTGGGGAGTACGGTCCGAAGATTAAACTCAAGGAATTGACGGGGACCC 910
QY 901 GCAACAGCGGTGATTTATGTAGATTAATTCGATGCAACGCGAAACCTTACCTACCTT 960
Db 911 GCAACAGCGGTGATTTATGTAGATTAATTCGATGCAACGCGAAACCTTACCTACCTT 970
QY 961 GACATGTAGCAATTTCTAGAGATAGATTAGTCT--TCGGAGACGTTAACAAGCTG 1017
Db 971 GACATGTAGCAATTTCTAGAGATAGATTAGTCTGCAAGAGAACTTAAACAAGCTG 1030
QY 1018 CTGCAATGCTGTGTGTCAGCTCTGTCTGTGAGATGTTGGTTAACTCCCGAGCAAGAGGCA 1077
Db 1031 CTGCAATGCTGTGTGTCAGCTCTGTCTGTGAGATGTTGGTTAACTCCCGAGCAAGAGGCA 1090
QY 1078 ACCCTTGCATTAATTCGCAATCATTTGTTGGGCACTTTAATGAGACTGCGCGGTGACAA 1137
Db 1091 ACCCTTGCATTAATTCGCAATCATTTGTTGGGCACTTTAATGAGACTGCGCGGTGACAA 1150
QY 1138 CCGAGAGAGGTGGAGTGAAGTCAAGTCTCATGAGCCCTTATGGGTAAGGCTTCAACAG 1197
Db 1151 CCGAGAGAGGTGGAGTGAAGTCAAGTCTCATGAGCCCTTATGGGTAAGGCTTCAACAG 1210
QY 1198 TAATACATGCGCGGTACAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAAGC 1257
Db 1211 TAATACATGCGCGGTACAGAGGTTGCCAACCTGCGAGAGGGAGCTAATCTCAGAAAGC 1270
QY 1258 GCGTCGATAGTCGCGATGCGAGTCTGCAACTGCACTCCGTGAAGTTCGGAATCGTAGTAAT 1317
Db 1271 GCGTCGATAGTCGCGATGCGAGTCTGCAACTGCACTCCGTGAAGTTCGGAATCGTAGTAAT 1330
QY 1318 CGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTACACACCGCGCTCAAC 1377
Db 1331 CGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTACACACCGCGCTCAAC 1390
QY 1378 CATGGAGTGGGTTTACACAGAGAGCTTAACTTAACCTTAAGAGAGGCGGTTTCCACAG 1437
Db 1391 CATGGAGTGGGTTTACACAGAGAGCTTAACTTAACCTTAAGAGAGGCGGTTTCCACAG 1449
QY 1438 TGAGATTCAATGAC 1450
Db 1450 TGAGATTCAATGAC 1462

```

```

RESULT 45
AY856079 1495 bp DNA linear BCT 12-JAN-2005
LOCUS Nitrosospira sp. En13 16S ribosomal RNA gene, partial sequence.
DEFINITION AY856079
ACCESSION AY856079.1 GI:57232110
KEYWORDS

```

```

SOURCE Nitrosospira sp. En13
ORGANISM Nitrosospira sp. En13
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosospira.
REFERENCE 1 (bases 1 to 1495)
AUTHORS Shaw,L.J., Nicol,G.W., Smith,Z., Fear,J., Prosser,J.I. and
Baggs,E.M.
TITLE Nitrous oxide production and nitrifier denitrification by
betaproteobacterial ammonia oxidising bacteria
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1495)
AUTHORS Shaw,L.J., Nicol,G.W., Smith,Z., Fear,J., Prosser,J.I. and
Baggs,E.M.
TITLE Direct Submission
JOURNAL Submitted (14-DEC-2004) School of Biological Sciences, University
of Aberdeen, Crutwickshank Building, St Machar Drive, Aberdeen AB24
3UU, U.K.

FEATURES
source Location/Qualifiers
1..1495 /organism="Nitrosospira sp. En13"
/mol_type="genomic DNA"
/strain="En13"
/db_xref="taxon:308023"
<1..>1495
/product="16S ribosomal RNA"

ORIGIN
Query Match 88.6%; Score 1290.8; DB 1; Length 1495;
Best Local Similarity 93.7%; Pred. No. 3.1e-13;
Matches 1368; Conservative 0; Mismatches 87; Indels 5; Gaps 2;

QY 1 ATTGAACGCTGCGCGCATGCTTTTACACATGCAAGTGAACGCGACGACGATCTTCAT 60
Db 1 ATTGAACGCTGCGCGCATGCTTTTACACATGCAAGTGAACGCGACGACGATCTTCAT 58
QY 61 CTGCTGGCGAGTGGCGGAGCGGTGAGTAATGCAATCGGAACGTATCCAGAAAGAGGGGATA 120
Db 59 CTGCTGGCGAGTGGCGGAGCGGTGAGTAATGCAATCGGAACGTATCTTCAGTGGGGATA 118
QY 121 ACCCATCAAGAAAGATGTGCTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGATCGAA 180
Db 119 ACCCATCAAGAAAGATGTGCTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGATCGAA 178
QY 181 GACCTTGCCTTTTGAAGCGCGCGATGTCGATTAGCTAATGCTGGGCTTAAAGCGCTAC 240
Db 179 GACCTTGCCTTTTGAAGCGCGCGATGTCGATTAGCTAATGCTGGGCTTAAAGCGCTAC 238
QY 241 CAAGCGACGATCAGTAGTGTCTGAGAGAGACGACGACACATGAGACTGAGACAGC 300
Db 239 CAAGCGACGATCAGTAGTGTCTGAGAGAGACGACGACACATGAGACTGAGACAGC 298
QY 301 GCCCAGACTCTTACGCGAGGACAGACAGTGGGAAATTTTGGCAATGGCGCAGACCTGATC 360
Db 299 GCCCAGACTCTTACGCGAGGACAGACAGTGGGAAATTTTGGCAATGGCGCAGACCTGATC 358
QY 361 CAGCAATGCCGCTGAGTGAAGAAAGCTTGGGTTTAAAGCTTTTCACTGAGAGAGA 420
Db 359 CAGCAATGCCGCTGAGTGAAGAAAGCTTGGGTTTAAAGCTTTTCACTGAGAGAGA 418
QY 421 AAAGGTACGTTAATTAATGTGACCCATGACGATGACAGAGAAAGACCGGGCTAAC 480
Db 419 AAAGGTACGTTAATTAATGTGACCCGTGACTAGACGATGACAGAGAAAGACCGGGCTAAC 478
QY 481 TAGTGCACAGACCGCGGTAAATTAAGTGAAGTGAATCCCGGGCTTAACTTGGCGT 540
Db 479 TAGTGCACAGACCGCGGTAAATTAAGTGAAGTGAATCCCGGGCTTAACTTGGCGT 538
QY 541 AAAGGTGCGACGCGCGCTTGTAGTCAATGTGAATCCCGGGCTTAACTTGGGAAT 600
Db 539 AAAGGTGCGACGCGCGCTTGTGTATGATGATGAAATCCCGGGCTTAACTTGGGAAT 598
QY 601 TGGGTTTGAATCAAAAGCTAGATGTGCGACAGGGAGGTGGAATTTCCATGTGTACAG 660

```

Db 599 TGGCTTTGAACACTAGAGCTAGAGTGTGGCAGAGGGGGGTGGATTCCAGCTGTAGCAG 658  
Qy 661 TGAATATGCTTATGATATGTAAGAAACATCGATGGCCGAGACCTCTCTGGGTTAACT 720  
Db 659 TGAATATGCTTATGATATGTAAGAAACATCGATGGCCGAGACCTCTCTGGGTTAACT 718  
Qy 721 GACGCTCATGACGAAAGCGTGGGGAGCAACAGATATGATACCTCTGGTATGCCAGCC 780  
Db 719 GACGCTCATGACGAAAGCGTGGGGAGCAACAGATATGATACCTCTGGTATGCCAGCC 778  
Qy 781 CTAAACGATGTCACTAGTTGTTGGCTTTATTAGGCTTGTATACGAAAGCTAAACGCTGA 840  
Db 779 CTAAACGATGTCACTAGTTGTTGGCTTTATTAGGCTTGTATACGAAAGCTAAACGCTGA 838  
Qy 841 AGTTGACCGCTGGGGAGTACCGTCCGAAAGATTAAATCTCAAGAAATGACGGGGACCC 900  
Db 839 AGTTGACCGCTGGGGAGTACCGTCCGAAAGATTAAATCTCAAGAAATGACGGGGACCC 898  
Qy 901 GCACAGCGGTGATATATGATGATTAATTCGATGCAACGCAAAACCTTACCTACCTT 960  
Db 899 GCACAGCGGTGATATATGATGATTAATTCGATGCAACGCAAAACCTTACCTACCTT 958  
Qy 961 GACATGTACGAAATTTTCTAGAGATGATTAGTGC--TTGGGAAACGCTTAAACAGAGTG 1017  
Db 959 GACATGTACGAAACCCGCTGAGAGGTGGGTGTCGCCGAAAGGAGCGGTTAACACAGTG 1018  
Qy 1018 CTGCAAGGCTGTCTGATGCTGCTGTGTGATAGATGTTGGGTTATGTCGCCGAAACAGAGCGA 1077  
Db 1019 CTGCAAGGCTGTCTGATGCTGCTGTGTGATAGATGTTGGGTTATGTCGCCGAAACAGAGCGA 1078  
Qy 1078 ACCCTGTGATTAATGATGATCACTTGTGGGCACTTTAATGAGACTCCGGTGAACAA 1137  
Db 1079 ACCCTGTGATTAATGATGATCACTTGTGGGCACTTTAATGAGACTCCGGTGAACAA 1138  
Qy 1138 CCGAGGAAGGTGGGGATGACGTCAAGTCTCATGGCCCTTATGGGTAGGGCTTCAACAG 1197  
Db 1139 CCGAGGAAGGTGGGGATGACGTCAAGTCTCATGGCCCTTATGGGTAGGGCTTCAACAG 1198  
Qy 1198 TAAATCAATGGCGCGTACAGAGGGTGGCCACCCCGGAGGGGAGCTAATTCAGAAAC 1257  
Db 1199 TAAATCAATGGCGCGTACAGAGGGTGGCCACCCCGGAGGGGAGCTAATTCAGAAAC 1258  
Qy 1258 GCGTGTAGTCCGATCGAGTCTGCACTGCACTCCGTAAGTCCGAATCCGTAGTAAT 1317  
Db 1259 GCGTGTAGTCCGATCGAGTCTGCACTGCACTCCGTAAGTCCGAATCCGTAGTAAT 1318  
Qy 1318 CCGGATCAAGATGTCGGGTGAATACGTTCCGGGTCTTGTATACACACCGCCGTCACAC 1377  
Db 1319 CCGGATCAAGATGTCGGGTGAATACGTTCCGGGTCTTGTATACACACCGCCGTCACAC 1378  
Qy 1378 CATGGAGTGGGTTTCAACGAAAGCAGTATGTTAAACGTTAAGAGAGGGCGCTTGGCAGG 1437  
Db 1379 CATGGAGTGGGTTTCAACGAAAGCAGTATGTTAAACGTTAAGAGAGGGCGCTTGGCAGG 1438  
Qy 1438 TGAGATTCACTGAGGAGTG 1457  
Db 1439 TGAGATTCACTGAGGAGTG 1458

RESULT 46  
NS16SR 1474 bp DNA linear BCT 23-FEB-1995  
LOCUS NS16SR  
DEFINITION Nitrosoespira briensis (strain C-128) 16S ribosomal RNA (16S rRNA)  
gene.  
ACCESSION L35505.1 GI:530899  
VERSION 16S ribosomal RNA; ribosomal RNA; small subunit.  
KEYWORDS Nitrosoespira briensis  
SOURCE Nitrosoespira briensis  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira.  
REFERENCE 1 (bases 1 to 1474)  
Tesse, A., Alm, E., Regan, J. M., Toze, S., Rittmann, B. E. and Stahl, D. A.

TITLE Evolutionary relationships among ammonia- and nitrite-oxidizing bacteria  
JOURNAL J. Bacteriol. 176 (21), 6623-6630 (1994)  
COMMENT 7961414  
FEATURES  
source  
1..1474  
/organism="Nitrosoespira briensis"  
/mol\_type="genomic DNA"  
/strain="C-128"  
/db\_xref="taxon:35799"  
1..1474  
/gene="16S rRNA"  
1..1474  
/gene="16S rRNA"  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 88.6%; Score 1290.6; DB 1; Length 1474;  
Best Local Similarity 93.6%; Pred. No. 3.2e-13;  
Matches 1366; Conservative 2; Mismatches 87; Indels 5; Gaps 2;  
Qy 1 ATTGAACCTGCGGCGATGCTTTACATGCAAGTCGAACGGCAGCAGGATGCTTGCAAT 60  
Db 17 ATTGAACCTGCGGCGATGCTTTACATGCAAGTCGAACGGCAGCAGGATGCTTGCAAT 74  
Qy 61 CTGGTGGGAGTGGGGAGCGGCTGATATGATGCAACGATACCAAGAGGGGGGAT 120  
Db 75 CTGGTGGGAGTGGGGAGCGGCTGATATGATGCAACGATACCAAGAGGGGGGAT 134  
Qy 121 ACGCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAAAGAGGAGATCGAAA 180  
Db 135 ACGCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAAAGAGGAGATCGCAA 194  
Qy 181 GACCTTGCGCTTTTGAGACCGCGCGATGCTGATTAAGTATGTTGGGTAAAGGCTTAC 240  
Db 195 GACCTTGCGCTTTTGAGACCGCGCGATGCTGATTAAGTATGTTGGGTAAAGGCTTAC 254  
Qy 241 CAAGCGCAAGATCACTAGTTGCTGAGAGAGCAACGCAACCTGGGACTGAGACAG 300  
Db 255 CAAGCGCTTGATCACTAGTTGCTGAGAGAGCAACGCAACCTGGGACTGAGACAG 314  
Qy 301 GCCCAGCTCCTCAACGGGAGGAGCAGTGGGGGATTTTGAACATGGGCGCAAGCTGATC 360  
Db 315 GCCCAGCTCCTCAACGGGAGGAGCAGTGGGGGATTTTGAACATGGGCGCAAGCTGATC 374  
Qy 361 CAGCAATCCCGGTGATGGAAGAGGCTTCGGGTGTTAAAGCTCTTTCAGTCGAAAGA 420  
Db 375 CAGCAATCCCGGTGATGGAAGAGGCTTCGGGTGTTAAAGCTCTTTCAGTCGAAAGA 434  
Qy 421 AAAGGTTACGTTAATATCCGATGCCATGACCGTATGACAGAAAGAACCCGCTAAC 480  
Db 435 AAAGGTTACGTTAATATCCGATGCCATGACCGTATGACAGAAAGAACCCGCTAAC 494  
Qy 481 TACGTGCCAGACGCCCGGTAAATACGTAGGGTCAAGGCTTAAATCGAATTAATCTGGGCT 540  
Db 495 TACGTGCCAGACGCCCGGTAAATACGTAGGGTCAAGGCTTAAATCGAATTAATCTGGGCT 554  
Qy 541 AAAGGTTCCGACGCGGCTTGTAAAGTGTAAATCCCGGGCTTAACTGGGAAT 600  
Db 555 AAAGGTTCCGACGCGGCTTGTAAAGTGTAAATCCCGGGCTTAACTGGGAAT 614  
Qy 601 TGGCTTTGAAATCTAAGAGCTTAGAGTGGCAGAGGGAGTGGGAATTCAGATGTAGAG 660  
Db 615 TGGCTTTGAAATCTAAGAGCTTAGAGTGGCAGAGGGAGTGGGAATTCAGATGTAGAG 674  
Qy 661 TGAATATGCTATGATATGTAAGAAACATCGATGGCCGAGACCTCTCTGGGTTAACT 720  
Db 675 TGAATATGCTATGATATGTAAGAAACATCGATGGCCGAGACCTCTCTGGGTTAACT 734  
Qy 721 GACGCTCATGACGAAAGCGTGGGGAGCAACAGATATGATACCTCTGGTATGCCAGCC 780  
Db 735 GACGCTCATGACGAAAGCGTGGGGAGCAACAGATATGATACCTCTGGTATGCCAGCC 794

QY 781 CTTAAACGATGTCATAGTGTGTGGCCCTTATTAGGCTTGTAACGAGTAAACGGCTGA 840  
Db 795 CTTAAACGATGTCATAGTGTGTGGCCCTTAAAGGGCTTGTAACGAGTAAACGGCTGA 854  
QY 841 AGTTGACCGCTGGGGAGTACGGTCCGAGATTTAAACTCAAGAAATGACGGGGACCC 900  
Db 855 AGTTGACCGCTGGGGAGTACGGTCCGAGATTTAAACTCAAGAAATGACGGGGACCC 914  
QY 901 GCACAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAGAACTTACCTACCTT 960  
Db 915 GCACAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAGAACTTACCTACCTT 974  
QY 961 GACATGATCGGAATTTTCTAGATAGATAGTAGTGC--TTCCGGAAACGCTTACACAGGTG 1017  
Db 975 GACATGATCGGAATTTTCTAGATAGATAGTAGTGC--TTCCGGAAACGCTTACACAGGTG 1034  
QY 1018 CTGCAATGCTGTGCTGACAGTCTGTGCTGATGATGTTGGTTAAAGTCCCGACAGAGCGCA 1077  
Db 1035 CTGCAATGCTGTGCTGACAGTCTGTGCTGATGATGTTGGTTAAAGTCCCGACAGAGCGCA 1094  
QY 1078 ACCCTGTCTATTATTTGCTCAATTTGGTGGGCACTTTATGATGACGCGGTGACAA 1137  
Db 1095 ACCCTGTCTATTATTTGCTCAATTTGGTGGGCACTTTATGATGACGCGGTGACAA 1154  
QY 1138 CCGGAGGAGGTGGGGAGTACGATCAAGTCCCTCAATGCGCTTATGGGTAGGGCTTCAACG 1197  
Db 1155 CCGGAGGAGGTGGGGAGTACGATCAAGTCCCTCAATGCGCTTATGGGTAGGGCTTCAACG 1214  
QY 1198 TAAATCAATGCGCGGTACAGAGGGTTGCCAACCCCGAGGGGGAGCTTAATCTCAAGAAC 1257  
Db 1215 TAAATCAATGCGCGGTACAGAGGGTTGCCAACCCCGAGGGGGAGCTTAATCTCAAGAAC 1274  
QY 1258 GCGTCTGATGTCGGGATGCGAGTCTGCAACTCGCTGAAGTGGGAATCGCTAGTAAT 1317  
Db 1275 GCGTCTGATGTCGGGATGCGAGTCTGCAACTCGCTGAAGTGGGAATCGCTAGTAAT 1334  
QY 1318 CCGGATCAGCATGTCGCGGTGAATTAAGTCCCGGGTCTTGTATACACACCGCCCTGCAC 1377  
Db 1335 CCGGATCAGCATGTCGCGGTGAATTAAGTCCCGGGTCTTGTATACACACCGCCCTGCAC 1394  
QY 1378 CATTGGAGTGGGTTTCAACAGAGCAGTACTTAAACCTTAAGAGGGCGCTTGCACCG 1437  
Db 1395 CATTGGAGTGGGTTTCAACAGAGCAGTACTTAAACCTTAAGAGGGCGCTTGCACCG 1454  
QY 1438 TGAGATTCTAGTACTGGGGTG 1457  
Db 1455 TGAGATTCTAGTACTGGGGTG 1474

RESULT 47  
AY684260 1458 bp DNA linear BCT 28-DEC-2004  
LOCUS Nitrosoospira sp. GS832 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY684260  
ACCESSION AY684260  
VERSION AY684260.1 GI:51012456  
KEYWORDS Nitrosoospira sp. GS832  
SOURCE Nitrosoospira sp. GS832  
ORGANISM Nitrosoospira sp. GS832  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosoospira.  
REFERENCE 1 (bases 1 to 1458)  
AUTHORS Ida, T., Satoh, M., Yabe, R., Takahashi, R. and Tokuyama, T.  
TITLE Identification of Genus Nitrosovibrio, Ammonia-Oxidizing Bacteria,  
by Comparison of N-Terminal Amino Acid Sequences of  
Phosphoglycerate Kinase  
J. Biochem. Bioeng. 98 (5), 380-383 (2004)  
2 (bases 1 to 1458)  
AUTHORS Takahashi, R., Yabe, R. and Tokuyama, T.  
TITLE Direct Submission  
JOURNAL Submitted (12-JUL-2004) Agricultural and Biological Chemistry,  
College of Bioresource Sciences, Nihon University, 1866 Kameino,  
Fujisawa, Kanagawa 252-8510, Japan

FEATURES  
source location/Qualifiers  
1..1458  
/organism="Nitrosoospira sp. GS832"  
/mol\_type="genomic DNA"  
/strain="GS832"  
/db\_xref="taxon:289297"  
<1..>1458  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 88.5%; Score 1289.2; DB 1; Length 1458;  
Best Local Similarity 93.6%; Pred. No. 3.4e-13;  
Matches 1367; Conservative 0; Mismatches 88; Indels 5; Gaps 2;

QY 1 ATTGAACGCTGGCGGCATCTTTTACATGCAAGTGAACCGGACAGACGGATGCTTCAT 60  
Db 1 ATTGAACGCTGGCGGCATCTTTTACATGCAAGTGAACCGGACAGACGGGGAAC-- 58  
QY 61 CTGCTGGCGAGTGGCGGACGGGTGATATGATCGGAACGTATCCAGAAAGGGGGGTA 120  
Db 59 CTGCTGGCGAGTGGCGGACGGGTGATATGATCGGAACGTATCTTCAAGTGGGGGAT 118  
QY 121 ACGCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAACAGGGGATGAAA 180  
Db 119 ACGCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAAAGAGGGGATGCA 178  
QY 181 GACCTTGGCTTTTGGAGCGGCCGATGTCTGATTAGCTAGTGGGTAAAGGCTTAC 240  
Db 179 GACCTTGGCTTTTGGAGCGGCCGATGTCTGATTAGCTAGTGGGTAAAGGCTTAC 238  
QY 241 CAAGGCGAGATACAGTATGATGCTGAGAGGACGACAGCAGCAGTGGGACTGAGACAG 300  
Db 239 CAAGGCTTGAATGATGATGCTGCTGAGAGGACGACAGCAGCAGTGGGACTGAGACAG 298  
QY 301 GCCCAGACTCTTACGGGAGGACAGCAGTGGGGAATTTTGAACATGGGCGCAAGCTTATC 360  
Db 299 GCCCAGACTCTTACGGGAGGACAGCAGTGGGGAATTTTGAACATGGGCGCAAGCTTATC 358  
QY 361 CAGCAATGCGCGGTGATGAGAAAGCCTTGGGTTGTAAGCTTTTCAGTGAAGAA 420  
Db 359 CAGCAATGCGCGGTGATGAGAAAGCCTTGGGTTGTAAGCTTTTCAGTGAAGAA 418  
QY 421 AAAGGTACCGTAAATTAATGTGACCCATGACGGTATCGACAGAAAGACCGGCTTAC 480  
Db 419 AAAGGTACCGTAAATTAATGTGACCCATGACGGTATCGACAGAAAGACCGGCTTAC 478  
QY 481 TAGGTGCAGCAGCGCGGTATATAGTGGGTCAAGCGTTAATCGAATTTACTGGGCGT 540  
Db 479 TAGGTGCAGCAGCGCGGTATATAGTGGGTCAAGCGTTAATCGAATTTACTGGGCGT 538  
QY 541 AAAGGTGCGCAGCGCGCTTGTAGTCAAGTGAATCCCGGGCTTAACCTGGGAAT 600  
Db 539 AAAGGTGCGCAGCGCGCTTGTAGTCAAGTGAATCCCGGGCTTAACCTGGGAAT 598  
QY 601 TGGCTTTGAAACTTAACAAGCTAAGTGTGCAAGGAGGTGGAATTCATGTGTACAG 660  
Db 599 TGGCTTTGAAACTTAACAAGCTAAGTGTGCAAGGAGGTGGAATTCATGTGTACAG 658  
QY 661 TGAATGCTGATGATATGGAAGAACATCGATGGCGAAGCAGCTCTGCGTTTAACT 720  
Db 659 TGAATGCTGATGATATGGAAGAACATCGATGGCGAAGCAGCTCTGCGTTTAACT 718  
QY 721 GACGCTGATGACGAAAGCGTGGGAGCAACAGATTAATACCTTGTGATGACAGCC 780  
Db 719 GACGCTGATGACGAAAGCGTGGGAGCAACAGATTAATACCTTGTGATGACAGCC 778  
QY 781 CTTAAACGATGTCATAGTGTGTGGGCTTATTAAGCTTGTATACGAAGCTTAACGCTGA 840  
Db 779 CTTAAACGATGTCATAGTGTGTGGGCTTATTAAGCTTGTATACGAAGCTTAACGCTGA 838  
QY 841 AGTTGACCGCTGGGGAGTACGGTCCGAGATTTAAACTCAAGAAATTAAGCGGGACCC 900  
Db 839 AGTTGACCGCTGGGGAGTACGGTCCGAGATTTAAACTCAAGAAATTAAGCGGGACCC 898

Qy	901	GCACAAGCGGGATTAATGTGGATTAATTCAGATGCAGCGCAAAAACCTTACCTT	960
Dd	899	GCAACAAGCGGGATTAATGTGGATTAATTCAGATGCAGCGCAAAAACCTTACCTT	958
Qy	961	GACATGTAGCGAAATTTCTAGAGATAGATTAGTGC--TTCCGGAACGCTAACACAGGT	1017
Dd	959	GACATGTATCCGAAGTCCGCTGTAGAGATGTGATGTGCCCCGAAGGGAACGTTAACACAGGT	1018
Qy	1018	CTGCATGGCTGTCCGTACACTCGTGTCTGTGAATGTCTGGGTTAACTCCCGCAACAGCGCA	1077
Dd	1019	CTGCATGGCTGTCCGTACACTCGTGTCTGTGAATGTCTGGGTTAACTCCCGCAACAGCGCA	1078
Qy	1078	ACCCCTTGCAATTAATTTGGCATTATTTGGTGGGCACTTAAATGAGACTGCCGTGACAA	1137
Dd	1079	ACCCCTTGCAATTAATTTGGCATTATTTGGTGGGCACTTAAATGAGACTGCCGTGACAA	1138
Qy	1138	CCGAGGAAGGTGGGAGTAGACGTCAAGTCTTCATAGGCCCTTAATGGGATGGGCTTACACG	1197
Dd	1139	CCGAGGAAGGTGGGAGTAGACGTCAAGTCTTCATAGGCCCTTAATGGGATGGGCTTACACG	1198
Qy	1198	TAAATCAATGGCGCTTACAGAGGTTGGCCAAACCCCGAGGGGAGCTAATCTCAGAAAGC	1257
Dd	1199	TAAATCAATGGCGCTTACAGAGGTTGGCCAAACCCCGAGGGGAGCTAATCTCAGAAAGC	1258
Qy	1258	GGCTGTAGTCCCGAATCCGAGCTCTCACTCGACTCCGCTGAAGTCCGAATCGTAGTAT	1317
Dd	1259	GGCTGTAGTCCCGAATCCGAGCTCTCACTCGACTCCGCTGAAGTCCGAATCGTAGTAT	1318
Qy	1318	CGCGGATCAGCATGTGCGGGTGAAATAGCTTCCGGGCTTTGTACACACGCCCGCTCACAC	1377
Dd	1319	CGCGGATCAGCATGTGCGGGTGAAATAGCTTCCGGGCTTTGTACACACGCCCGCTCACAC	1378
Qy	1378	CATGGAAATGGGTTTACACAGAGAGGTAAGTTAACCGTAAAGAGGGCGCTTGCCACCG	1437
Dd	1379	CATGGAAATGGGTTTACACAGAGAGGTAAGTTAACCGTAAAGAGGGCGCTTGCCACCG	1438
Qy	1438	TGAGATTTCATGACTGGGGTG 1457	
Dd	1439	TGAGATTTCATGACTGGGGTG 1458	
RESULT 48			
LOCUS	AY123790	1497 bp	DNA linear BCT 23-SEP-2003)
DEFINITION	Nitrosoespira sp. Nsp62 16S ribosomal RNA gene, partial sequence.		
ACCESSION	AY123790		
VERSION	AY123790.1	GI:24474419	
KEYWORDS			
SOURCE	Nitrosoespira sp. Nsp62		
ORGANISM	Nitrosoespira sp. Nsp62		
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;		
AUTHORS	Nitrosomonadaceae; Nitrosoespira.		
TITLE	1 (bases 1 to 1497)		
	Pukhkhod, U., Wagner, M., Timmermann, G., Pommerening-Rosser, A. and		
JOURNAL	Koops, H.P.		
	16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial		
PUBMED	ammonia-oxidizing isolates: extension of the dataset and proposal		
REFERENCE	of a new lineage within the nitrosomonads		
AUTHORS	Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)		
TITLE	13130037		
	2 (bases 1 to 1497)		
JOURNAL	Pukhkhod, U., Timmermann, G., Koops, H.-P., Pommerening-Rosser, A. and		
PUBMED	Wagner, M.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische		
FEATURES	Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany		
	Location/Qualifiers		
source	1. 1497		
	/organism="Nitrosoespira sp. Nsp62"		
	/mol_type="genomic DNA"		
	/strain="Nsp62"		
	/db_xref="taxon:200117"		

ORIGIN	Query Match	88.5%;	Score 1289.2;	DB 1;	Length 1497;
	Beat Local Similarity	93.6%;	Pred. No. 3.3e-13;		
	Matches 1367;	Conservative 0;	Mismatches 88;	Indels 5;	Gaps 2
ORNA	<1..>.1497				
	/product="16S ribosomal RNA"				
QY	1 ATTGAAGCCTGGCCGACATGCTTTACATGCAAGTCGAAACGGCACACAGGAGTGGCTTGCAT	60			
DB	3 ATTGAAGCCTGGCCGACATGCTTTACATGCAAGTCGAAACGGCACACAGGAGGCAAC--	60			
QY	61 CTGGTGGGAGATGGGCGGACGGGTGATGATCATCGAAGCATCAGAAAGAGGGGGGATA	120			
DB	61 CTGGTGGGAGATGGGCGGACGGGTGATGATCATCGAAGCATCTTCAAGTGGGGGATA	120			
QY	121 ACGCATCGAAGATGTGTCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAA	180			
DB	121 ACGCAGCGAAGATGTGTCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGCAA	180			
QY	181 GACCTTGGGCTTTTGGAGCGGCGCATGTCTGATTAGCTAAGTTGGTGGGGTAAAGGCTTAC	240			
DB	181 GACCTTGGGCTTTTGGAGCGGCGCATGTCTGATTAGCTAAGTTGGTGGAGTAAAGGCTTAC	240			
QY	241 CAAGGCGACATCATGATGATGTGTCTGAAAGGACGACACGACCATCTGGGATCTGAGACAG	300			
DB	241 CAAGGCTTGCATCATGATGATGTGTCTGAAAGGACGACACGACCATCTGGGATCTGAGACAG	300			
QY	301 GCCGACATCTCTTAACGGGAGGACGACGATGGGGAAATTTTGGACATATGGGCGCAAGCCTGATC	360			
DB	301 GCCGACATCTCTTAACGGGAGGACGACGATGGGGAAATTTTGGACATATGGGCGCAACCTGATC	360			
QY	361 CAGCATGCCGCGGTGATGAAAGAGGCTTGGGGTTGTAAGCTTTTCACTCGAAGAGA	420			
DB	361 CAGCATGCCGCGGTGATGAAAGAGGCTTGGGGTTGTAAGCTTTTCAAGCGGAGAGA	420			
QY	421 AAAGGTTACGGTAAATATATCGTGAACCCATGACGGATGACAGAAAGAACCCGGCTAAC	480			
DB	421 AAAGGTTACGGCTAATATATCGTGAACCCATGACGGATGACCGAAGAAAGAACCCGGCTAAC	480			
QY	481 TACGTGCACGACCGCGGTAAATACGTAGGGTGCAGAGGCTTAATCGAATTACTGGGCGT	540			
DB	481 TACGTGCACGACCGCGGTAAATACGTAGGGTGCAGAGGCTTAATCGAATTACTGGGCGT	540			
QY	541 AAAGGTTGCAGGCGGCTTTGTAAGTGAATGTGAATCCCGGGCTTAACTTGGGAAT	600			
DB	541 AAAGGTTGCAGGCGGCTTTGTAAGTGAATGTGAATCCCGGGCTTAACTTGGGAAC	600			
QY	601 TGGCTTTGAAACTTAAGTGAAGTGTGGAGGAGGAGGGTGGAAATTCATGTGTAGAGAG	660			
DB	601 TGGCTTTGAAACTTAAGTGAAGTGTGGAGGAGGAGGGTGGAAATTCATGTGTAGAGAG	660			
QY	661 TGAATATGGTATGATATATGAAGAACATCGATGGCGAAGGACGCTCTCGGGTTAACT	720			
DB	661 TGAATATGGTATGATATGATGAAGAACATCGATGGCGAAGGACGCTCTCGGGTTAACT	720			
QY	721 GAGCCTCATGACGAAAGCGTGGGAGCAACAGGATTAAGTACCCTGTGTGTCCAGGCC	780			
DB	721 GAGCCTCATGACGAAAGCGTGGGAGCAACAGGATTAAGTACCCTGTGTGTCCAGGCC	780			
QY	781 CTAAACGATGTCATCTAGTTTGTGGGCTTATTAAGCTTTGGATTAACGAAGCTTAAGCGCGTA	840			
DB	781 CTAAACGATGTCATCTAGTTTGTGGGCTTATTAAGCGGCTTGGATTAAGCGCGTA	840			
QY	841 AGTTGACCGCTGGGGAGTACGCTCGCAAGATTAAATCTCAAGGAAATTGACGGGACCC	900			
DB	841 AGTTGACCGCTGGGGAGTACGCTCGCAAGATTAAATCTCAAGGAAATTGACGGGACCC	900			
QY	901 GCACAAGCGGTGATTAATGTGGATTAAATTCGATGCACGCGAATAAACCCTTACCTT	960			
DB	901 GCACAAGCGGTGATTAATGTGGATTAAATTCGATGCACGCGAATAAACCCTTACCTT	960			
QY	961 GACATGTAGCGAATTTTCTAAGATTAAGTGTGCT--TTGGGAGACGCTTAACAGGTG	101			

Db	961	GACATGTATCCGAAACCCCGCTGAGAGTGAGGCGCTGCCGAAAGGAGCGGTAACTCAAGSTG	1020
Qy	1018	CTGCATGGCTGCTCGTCAAGCTGTCGTCGTGAGATGTTGGTTAACTCCGCAACGAGCGCA	1077
Db	1021	CTGCATGGCTGCTCGTCAAGCTGTCGTCGTGAGATGTTGGTTAACTCCGCAACGAGCGCA	1080
Qy	1078	ACCCTTGTCAATTAATTCGATCATATTGGTTGGGCACTTTATGAGACTGCCGGTGACAA	1137
Db	1081	ACCCTTGTCAATTAATTCGATCATATTGGTTGGGCACTTTATGAACTGCCGGTGACAA	1140
Qy	1138	CCGAGGAAGGTGGGGATGACGTCAAGCTCATGGCCCTTAATGAGGCTTCAACG	1197
Db	1141	CCGAGGAAGGTGGGGATGACGTCAAGCTCATGGCCCTTAATGAGGCTTCAACG	1200
Qy	1198	TAATCAATGGCGCGTCACAGAGGTTGGCAACCCGCGAGGGGAGCTTAATTCAGAAAGC	1257
Db	1201	TAATCAATGGCGCGTCACAGAGGTTGGCAACCCGCGAGGGGAGCCATTCAGAAAGC	1260
Qy	1258	GCGTGTGTAGTCCGGATTCGGAGTCTTGCAACTGATCCGTGAAGTCGAAATGCGTAGTAAT	1317
Db	1261	GCGTGTGTAGTCCGGATTCGGAGTCTTGCAACTGATCCGTGAAGTCGAAATGCGTAGTAAT	1320
Qy	1318	GCGGATCAGCATGTCCGGGTGAATATACCTTCCGGGTCTTGTACACACCGCCGCTCAC	1377
Db	1321	GCGGATCAGCATGTCCGGGTGAATATACCTTCCGGGTCTTGTACACACCGCCGCTCAC	1380
Qy	1378	CATGGAGTGGGTTTCACCAAGCAGGTAGTTAAACCGTAAGAGGCGCTTGCCACGG	1437
Db	1381	CATGGAGTGGGTTTCACCAAGCAGGTAGTTAAACCGCAAGAGGCGCTTGCCACGG	1440
Qy	1438	TGAGATTCAATGCTGGGGTG	1457
Db	1441	TGAGATTCAATGCTGGGGTG	1460

RESULT	49
AY123802	
LOCUS	1497 bp DNA linear ECT 23-SEP-2003
DEFINITION	Nitrosospirita sp. Nsp2 16S ribosomal RNA gene, partial sequence.
ACCESSION	AY123802
VERSION	AY123802.1 GI:24474431
KEYWORDS	.
SOURCE	Nitrosospira sp. Nsp2
ORGANISM	Nitrosospira sp. Nsp2 Bacteria; Proteobacteria; Betaproteobacteriales; Nitrospirales; Nitrosoarchaeota; Nitrospirochaetes;
REFERENCE	Nitrosomonadaceae; Nitrospira. 1 (bases 1 to 1497)
AUTHORS	Purkhof,U., Wagner,M., Timmermann,G., Pommerening-Rooser,A. and Koops,H.P.
TITLE	16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal of a new lineage within the nitrospirads
JOURNAL	Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003) PUBMED 13130037
REFERENCE	2 (bases 1 to 1497)
AUTHORS	Purkhof,U., Timmermann,G., Koops,H.-P., Pommerening-Roeser,A. and Wagner,M.
TITLE	Direct Submission
JOURNAL	Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany Location/Qualifiers
FEATURES	1..1497
source	/organism="Nitrosospira sp. Nsp2" /mol_type="genomic DNA" /strain="Nsp2" /db_xref="taxon:136548"  
rRNA	<1..>1497 /product="16S ribosomal RNA"
ORIGIN	
Query Match	88.5%; Score 1289.2; DB 1; Length 1497;
Best Local Similarity	93.6%; Pred. No. 3.3e-13;

	Matches	136/	Conservative	0;	Mismatches	88;	Indels	5;	Gaps	2;
QY	1	ATTGAACGCTGGCGGCGATGCTTTTACATGACAGTGCAGACGGCAGCACGGAGTCTTGCGAT	60							
Db	3	ATTGAACGCTGGCGGCGATGCTTTTACATGACAGTGCAGACGGCAGCACGGAGGCGCAAC--	60							
QY	61	CTGGTGGCGAGTGGCGGACGGGTGAGTTATGCATCGGAACGTATCCAGAAAGAGGGGGGTA	120							
Db	61	CTGGTGGCGAGTGGCGGACGGGTGAGTTATGCATCGGAACGTATCCAGAAAGAGGGGGGATA	120							
QY	121	AGCGATCGAAAGATGTGCTTAATCCGCATATATCTTAAGGAGAAAGACAGGGGATCGAAA	180							
Db	121	AGCGACCGAAAGGTGTGCTTAATCCGCATATATCTTCTGGAGAAAGACAGGGGATCGCAA	180							
QY	181	GACCTTGCGCTTTTGGAGCGGCCGATGTCTGATTAGCTAGTTGGTGGGGGTTAAAGCCCTAC	240							
Db	181	GACCTTGCGCTTTTGGAGCGGCCGATGTCTGATTAGCTAGTTGGTGGGGGTTAAAGGCTTAC	240							
QY	241	CAAGCGGACGATCAGTAGTTGGTCTTGAGAGACGACACGACCACTGAGGACTGAGACAG	300							
Db	241	CAAGCGGTGATCAGTAGTGTGTCTGAGAGAGACGACCGACCACTGGGACTGAGACAG	300							
QY	301	GCCCAAGCTCTTACCGGAGGACACATGGGGGAAATTTTGGCAATPGGCGGAGCTGATC	360							
Db	301	GCCCAAGCTCTTACCGGAGGACACATGGGGGAAATTTTGGCAATPGGGGAAACCTTATC	360							
QY	361	CAGCAATGCCGCGTGAAGTGAAGAAAGGCGCTTCGGGTGTGAAGCTCTTTCAGTCGAGAGA	420							
Db	361	CAGCCATGCCGCGTGAAGTGAAGAAAGGCGCTTCGGGTGTGAAGCTCTTTCAGCGGAGAGA	420							
QY	421	AAAGGTTACGGTAAATTAATGCTGAGCCCATGACGGTATCGACAGAGAGACACGGGCTAAC	480							
Db	421	AAAGGTTACGGGCTTAATATCCGTGACAACTGACGGTACCGAGAGAGAGACACCGGCTAAC	480							
QY	481	TAGTGCCAGACAGCCCGCGTAAATACGTAGGGGTGCAGACGCTTAATCCGGAATTACTGGGGGT	540							
Db	481	TAGCTGCGACACAGCCCGCGTAAATACGTAGGGGTGCAGACGCTTAATCCGGAATTACTGGGGGT	540							
QY	541	AAAGGGGCGCAGGCGGCGCTTGTAAATCAGATGTGAATCCCGGGCTTAAACCTGGGAAT	600							
Db	541	AAAGCGTGGCAGGCGGCTTGTAAATCAGATGTGAATCCCGGGCTTAAACCTGGGAAC	600							
QY	601	TGCGTTTGAATCTAACAAAGCTAGATGTGCGACAGGAGGTGCAATTCATGTGTAGCAG	660							
Db	601	TGCGTTTGAATCTAACAGGCTAGATGTGCGACAGGAGGGGTGGAAATTCACATGTGTACAG	660							
QY	661	TGAATATCGTGAAGATATGGAAGAAACATTCATGCGGAAGGACGCTCTGGGGTTAACACT	720							
Db	661	TGAATATCGTAGAATATGGAAGAAACATTCATGCGGAAGGACGCTCTGGGGTTAACACT	720							
QY	721	GAGCGTCATGACGAAAGCGTGGGGGACAAACAGATTTAGTATCCCTGTAGTCCAGGCC	780							
Db	721	GAGCGTCAGGACGAAAGCGTGGGGGACAAACAGATTTAGTATCCCTGTAGTCCAGGCC	780							
QY	781	CTAAACGATGCACTAGTTGTTGGGCTTATTAGGCTTGGTTACGAAAGCTAACGCGTGA	840							
Db	781	CTAAACGATGCACTAGTTGTTGGGCTTATTAGGCTTGGTTACGCAAGCTAACGCGTGA	840							
QY	841	AGTTGACCGCTGGGGGAGTACGCTGCGCAAGATTAAACTCAAGAAATTGACGGGGACCC	900							
Db	841	AGTTGACCGCTGGGGGAGTACGCTGCGCAAGATTAAACTCAAGAAATTGACGGGGACCC	900							
QY	901	GCAACAGGGGTGATTAATGTGGATTTAATTCGATGCAACGCGAATAAACCCTTACCTT	960							
Db	901	GCAACAGGGGTGATTAATGTGGATTTAATTCGATGCAACGCGAATAAACCCTTACCTT	960							
QY	961	GACATGTAGCGAAATTTTCTAGAGATAGATTAGTGC--TTGCGGAACGCTTAACACAGGTG	1017							
Db	961	GACATGTAGCGAAACCCCGGTAGAGGTGGGGGTGCCGAAAGGAGCGGTAAACACAGGTG	1028							
QY	1018	CTGCATGGCTGTGCTCAGCTGTGTGCTGAGATGTGGGTTAAATCCCGCAACGACGGCA	1077							
Db	1021	CTGCATGGCTGTGCTCAGCTGTGTGCTGAGATGTGGGTTAAAGTCCCGCAACGACGGCA	1086							

QY	1078	ACCCTTGTCATTAATTGCCATCATTTGGTGGGCACTTTAATGAGACTGCCGGTGACAA	1137
Db	1081	ACCTTGTCATTAATTTGGCATCATTCAGTTGGGCACTTTAATGAAACTGCCGGTGACAA	1140
QY	1138	CCGAGGAAGGTGGGGATGACGCTCAAGTCCTCATGGCCCTTATGGGTATGGGCTTCACACG	1197
Db	1141	CCGAGGAAGGTGGGGATGACGCTCAAGTCCTCATGGCCCTTATGGGTATGGGCTTCACACG	1200
QY	1198	TAATCAATGGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGGAGCTAATCTCAGAAAGC	1257
Db	1201	TAATCAATGGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGGAGCTAATCTCAGAAAGC	1260
QY	1258	GCGTCGTAGTCCCGATTCGGAGCTCTGCAACTGCACTCCGTGAAAGTCGGAATGGCTAGTAAT	1317
Db	1261	GCGTCGTAGTCCCGATTCGGAGCTCTGCAACTGCACTCCGTGAAAGTCGGAATGGCTAGTAAT	1320
QY	1318	CGCGGATCAGCATGTCCGCGGTGAATACGTTCCCGGGTCTTTGTACACACCGCCCGTCACAC	1377
Db	1321	CGCGGATCAGCATGTCCGCGGTGAATACGTTCCCGGGTCTTTGTACACACCGCCCGTCACAC	1380
QY	1378	CATGGAGTGGGTTTCACCAAGCAGGTAGTTAACCCTAAGAGAGGGCGCTTGCCACCG	1437
Db	1381	CATGGAGTGGGTTTCACCAAGCAGGTAGTTAACCCTAAGAGAGGGCGCTTGCCACCG	1440
QY	1438	TGAGATTCATGACTGGGGTG	1457
Db	1441	TGAGATTCATGACTGGGGTG	1460
RESULT 50			
LOCUS	AY123803	1497 bp	DNA linear BCT 23-SEP-2003
DEFINITION	Nitrosobrevibrio tenuis 16S ribosomal RNA gene, partial sequence.		
ACCESSION	AY123803		
VERSION	AY123803.1	GI:24474432	
KEYWORDS	Nitrosobrevibrio tenuis		
SOURCE	Nitrosobrevibrio tenuis		
ORGANISM	Bacteria; Proteobacteria; Betaproteobacteria; Nitrospomonadales; Nitrospomonadaceae; Nitrosobrevibrio.		
REFERENCE	1 (bases 1 to 1497)		
AUTHORS	Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Rosser, A. and Koops, H.-P.		
TITLE	16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal of a new lineage within the nitrospomonadales		
JOURNAL	Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)		
PUBMED	13130037		
REFERENCE	2 (bases 1 to 1497)		
AUTHORS	Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Rosser, A. and Wagner, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany		
FEATURES	Location/Qualifiers		
SOURCE	1..1497		
	/organism="Nitrosobrevibrio tenuis"		
	/mol_type="genomic DNA"		
	/strain="Nv1"		
	/db_xref="taxon:1233"		
	/note="type strain"		
rRNA	<1..>1497		
	/product="16S ribosomal RNA"		
ORIGIN			
Query Match			
Beat Local Similarity 88.5%; Score 1289.2; DB 1; Length 1497;			
Matches 1367; Conservative 0; Mismatches 88; Indels 5; Gaps 2;			
QY	1	ATTGAAGCGTGGCGGCGATGCTTTACATGCAAGTCGAACGGCAGACGAGTGGCTTGAT	60
Db	3	ATTGAAGCGTGGCGGCGATGCTTTACATGCAAGTCGAACGGCAGACGAGGCGCAAC--	60

QY	61	CTGGTGGGAGTGGGGGACCGGGTAGATTA	TGCAATCGAAACGTAATCCAAAGAGGGGGGTA	120
Db	61	CTGGTGGGAGTGGGGAACGGGGTAGATTA	TGCAATCGAAACGTAATCTTCGATGGGGGGATA	120
QY	121	ACGCATCCAAAGATGTGCTAATATACCAG	CAATATATCTTAAGGAGGAAAGCAGGGGATCGAAA	180
Db	121	AACGACCGAAAGGTGTGCTAATATACGG	CAATATCTTAACGGAAGAAAGCAGGGGATCGCAA	180
QY	181	GACCTTGGCGCTTTTGGAGCGGCCGAT	GTCTGATTAGCTAGTTGGTGGGGTAAAGGCTTAC	240
Db	181	GACCTTGGCGCTTCTTAGACGGCGCGAT	GTCTGATTAGCTAGTTGGTGGGGTAAAGGCTTAC	240
QY	241	CAAGGCGACGATCACTAGTTGGTCTGAG	AGGACGACACGACCAACTGCGGACCTGAGACACG	300
Db	241	CAAGGCGACGATCACTAGTTGGTCTGAG	AGGACGACGACACCACTGCGGACCTGAGACACG	300
QY	301	GCCCAAGCTCCTTACGGGAGGACGACAG	TGGGGAATTTTGGACAATGGGCGCAAGCCTGATC	360
Db	301	GCCCAAGCTCCTTACGGGAGGACGACAG	TGGGGAATTTTGGACAATGGGCGCAAGCCTGATC	360
QY	361	CAGCAATCCCGGTGATGATGAGAAGG	CGCTTGGTAAACCTCTTACGTACGACGAAGA	420
Db	361	CAGCCATCCCGGTGATGATGAGAAGG	CGCTTGGTAAACCTCTTACGTACGACGAAGA	420
QY	421	AAAGGTTACGGTAAATTAATCTGTGAC	CCATGACGATTCGACAGAAAGCACACCGGCTAAC	480
Db	421	AAAGGTTACGGGCTAATATCCCGTGA	CTACTGACGATACCGGAAGAAAGCACACCGGCTAAC	480
QY	481	TACGTGGCAGACGCGCGGGTAAATAC	TAGTGGGTGCAAGGCTTAATGCGAATTAATCTGGGCGT	540
Db	481	TACGTGGCAGACGCGCGGGTAAATAC	TAGTGGGTGCAAGGCTTAATGCGAATTAATCTGGGCGT	540
QY	541	AAAGGGTACGACGAGCGCGCTTGTAT	AGTCAAGTGTAAATCCCGGGCTTAACTTGGGAAT	600
Db	541	AAAGGGTACGACGAGCGCGCTTGTAT	AGTCAAGTGTAAATCCCGGGCTTAACTTGGGAAC	600
QY	601	TGCGTTTGAAACTACAAAGCTAAGAT	GTGGCGAGAGGTGAAATTCATGTGTAGCAG	660
Db	601	TGCGTTTGAAACTACAAAGCTAAGAT	GTGGCGAGAGGTGAAATTCACCTGTGTAGCAG	660
QY	661	TGAAATGGGTAGAGTATGGAAGAACAT	CGATGGGGAAGGCGAGCCTCCTGGGGTTAACACT	720
Db	661	TGAAATGGGTAGAGTATGGAAGAACAC	CGGATGGGGAAGGCGAGCCTCCTGGGGTTAACACT	720
QY	721	GACGCTACGACAGAAACGTTGGGAG	CAAAACAGGATTAAGATACCTGGTAGTCCACGCC	780
Db	721	GACGCTACGACAGAAACGTTGGGAG	CAAAACAGGATTAAGATACCTGGTAGTCCACGCC	780
QY	781	CTAAACGATGTCAACTAGTTGTTGGG	CCTTATTAGCTTTGGTATACGAAGCTAACCGGTGA	840
Db	781	CTAAACGATGTCAACTAGTTGTTGGG	CCTTATTAGCTTTGGTATACGAAGCTAACCGGTGA	840
QY	841	AGTTGACCGGCTGGGGGAGTACGGT	GTGGCAAGATTAATACTCAAGAGAAATGACGGGGACCC	900
Db	841	AGTTGACCGGCTGGGGGAGTACGGT	GTGGCAAGATTAATACTCAAGAGAAATGACGGGGACCC	900
QY	901	GCACAAGCGGTGATTAATGTGAGATTA	ATTCATGCAACGCGGAAAACTTACCTTACCTT	960
Db	901	GCACAAGCGGTGATTAATGTGAGATTA	ATTCATGCAACGCGGAAAACTTACCTTACCTT	960
QY	961	GACATGTAGCGAATTTTCTAGAGATNA	ATTAGTGC---TTGGGGAACGCTAACACAGGTG	1020
Db	961	GACATGTAGCGAATTTTCTAGAGATNA	ATTAGTGC---TTGGGGAACGCTAACACAGGTG	1020
QY	1018	CTGCATGGCTGTCTGACAGCTCGTGT	CGTAGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA	1077
Db	1021	CTGCATGGCTGTCTGACAGCTCGTGT	CGTAGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA	1080
QY	1078	ACCTTGTGATTAATGTCATCACTTTGG	TGGGCACTTTAATGAGACATGCGCGGTACAA	1133
Db	1081	ACCTTGTGATTAATGTCATCACTTTGG	TGGGCACTTTAATGAGACATGCGCGGTACAA	1144

QY 1138 CCGAGGAAAGTGGGATGACGTCAAGTCTCATGSCCTTATGGTAGGGCTTTCACAGC 1197  
|  
|  
|  
Db 1141 CCGAGGAAAGTGGGATGACGTCAAGTCTCATGSCCTTATGGTAGGGCTTTCACAGC 1200  
|  
|  
|  
QY 1198 TAATACAAATGGCCGCTTACAGAGGGTGTCCAAACCCGAGAGGGAGCTTAATCTCAGAAAGC 1257  
|  
|  
|  
Db 1201 TAATACAAATGGCCGCTTACAGAGGGTGTCCAAACCCGAGAGGGAGCTTAATCTCAGAAAGC 1260  
|  
|  
|  
QY 1258 GGGTGTAGTCCGAGTCCGAGTCTGCAACTGCACTCCGTAAGTCCGAAATCGCTAGTAAT 1317  
|  
|  
|  
Db 1261 GGGTGTAGTCCGAGTCCGAGTCTGCAACTGCACTCCGTAAGTCCGAAATCGCTAGTAAT 1320  
|  
|  
|  
QY 1318 CCGGATCAGCATGTGCGGGTGAATACGTTCCGGGTCTTTGTAACAACCGCCGCTCACAC 1377  
|  
|  
|  
Db 1321 CCGGATCAGCATGTGCGGGTGAATACGTTCCGGGTCTTTGTAACAACCGCCGCTCACAC 1380  
|  
|  
|  
QY 1378 CATGGAGTGGGTTTACACCAAGACAGTAGTCTAACCCGTAAGAGGGCGCTTGCCACCG 1437  
|  
|  
|  
Db 1381 CATGGAGTGGGTTTACACCAAGACAGTAGTCTAACCCGTAAGAGGGCGCTTGCCACCG 1440  
|  
|  
|  
QY 1438 TGAATTCATGACTGGGGTG 1457  
|  
|  
|  
Db 1441 TGAATTCATGACTGGGGTG 1460  
|  
|  
|  
RESULT 51  
AY123787 1497 bp DNA linear BCT 23-SEP-2003  
LOCUS Nitrosospiira sp. Nsp40 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY123787  
ACCESSION AY123787  
VERSION AY123787.1 GI:24474416  
KEYWORDS Nitrosospiira sp. Nsp40  
SOURCE Nitrosospiira sp. Nsp40  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospiira.  
REFERENCE 1 (bases 1 to 1497)  
AUTHORS Purkhof, U., Wagner, M., Timmermann, G., Pommerening-Rosser, A. and  
Koops, H. P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the Nitrosomonads  
JOURNAL Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
PUBMED 13130037  
REFERENCE 2 (bases 1 to 1497)  
AUTHORS Purkhof, U., Timmermann, G., Koops, H.-P., Pommerening-Rosser, A. and  
Wagner, M.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Leibniz-Institut fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1. 1497  
/organism="Nitrosospiira sp. Nsp40"  
/mol\_type="genomic DNA"  
/strain="Nsp40"  
/db\_xref="taxon:200114"  
<1..>1497  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 88.4%; Score 1287.6; DB 1; Length 1497;  
Beat Local Similarity 93.6%; Pred. No. 3.5e-13;  
Matches 1366; Conservative 0; Mismatches 89; Indels 5; Gaps 2;  
QY 1 ATGAAGCCTGCGCGCATGCTTTACATGCAATGTCGAACGCGACGACGAGTGTGTAT 60  
|  
|  
|  
Db 3 ATGAAGCCTGCGCGCATGCTTTACATGCAATGTCGAACGCGACGACGAGTGTGTAT 60  
|  
|  
|  
QY 61 CTGGTGGGAGTGGCGGAGCGGGTAGTAATGATCGAAGCGTATCCAGAAGAGGGGGGTA 120  
|  
|  
|  
Db 61 CTGGTGGGAGTGGCGGAGCGGGTAGTAATGATCGAAGCGTATCCAGAAGAGGGGGGTA 120  
|  
|  
|  
QY 121 ACGATCGAAGAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGCAGGGGATCGAA 180  
|  
|  
|

Db 121 ACGATCGAAGAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGCAGGGGATCGAA 180  
|  
|  
|  
QY 181 GACCTTGGCTTTTGAAGCGCCGATGTCTGATTAGCTAGTGGTGGGTTAAAGCCTTAC 240  
|  
|  
|  
Db 181 GACCTTGGCTTTTGAAGCGCCGATGTCTGATTAGCTAGTGGTGGGTTAAAGCCTTAC 240  
|  
|  
|  
QY 241 CAAGCGAAGATCAGTATGTTGGTCTGAGAGGACGACGCGCACATCGGGACTGAGACAG 300  
|  
|  
|  
Db 241 CAAGCGTTCAGTACAGTATGTTGGTCTGAGAGGACGACGCGCACATCGGGACTGAGACAG 300  
|  
|  
|  
QY 301 GCCCAAGTCTTACGGGAGGACAGATGGGGGAATTTTGGACAATGGGCGCAACCTGATC 360  
|  
|  
|  
Db 301 GCCCAAGTCTTACGGGAGGACAGATGGGGGAATTTTGGACAATGGGCGCAACCTGATC 360  
|  
|  
|  
QY 361 CAGCAATGCGCGTGAAGTGAAGAGCGCTTCCGGTTTAAAGCTCTTTCAGTGCAGAA 420  
|  
|  
|  
Db 361 CAGCAATGCGCGTGAAGTGAAGAGCGCTTCCGGTTTAAAGCTCTTTCAGTGCAGAA 420  
|  
|  
|  
QY 421 AAAGTTACGCTAATATATCTGTAACCCATGACGCTATCGACGAAGACACCGGCTAAC 480  
|  
|  
|  
Db 421 AAAGTTACGCTAATATATCTGTAACCCATGACGCTATCGACGAAGACACCGGCTAAC 480  
|  
|  
|  
QY 481 TACGTCCAGACGCGCGGTAATACGTAAGGTTGCAAGCGTTAATCGGAATTTACTGGCGT 540  
|  
|  
|  
Db 481 TACGTCCAGACGCGCGGTAATACGTAAGGTTGCAAGCGTTAATCGGAATTTACTGGCGT 540  
|  
|  
|  
QY 541 AAAGGTGCGCAGCGCGGCTTGTAGTCAAGTGTGAATCCCGGCTTAACCTGGGAAT 600  
|  
|  
|  
Db 541 AAAGGTGCGCAGCGCGGCTTGTAGTCAAGTGTGAATCCCGGCTTAACCTGGGAAT 600  
|  
|  
|  
QY 601 TGGCTTTGAACTACAAAGCTAGTAGTGTGGCAGAGGAGGTGAATTCATGTGTACAG 660  
|  
|  
|  
Db 601 TGGCTTTGAACTACAAAGCTAGTAGTGTGGCAGAGGAGGTGAATTCATGTGTACAG 660  
|  
|  
|  
QY 661 TGAATGCGTAAGATATGGAAGAACATGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
|  
|  
|  
Db 661 TGAATGCGTAAGATATGGAAGAACATGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
|  
|  
|  
QY 721 GACGCTATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTTAGTCCAGCC 780  
|  
|  
|  
Db 721 GACGCTATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTTAGTCCAGCC 780  
|  
|  
|  
QY 781 CTAAACGATGCTCACTAGTGTGGGCTTATTAAGCTTGGTAAAGCAAGCGGTGA 840  
|  
|  
|  
Db 781 CTAAACGATGCTCACTAGTGTGGGCTTATTAAGCTTGGTAAAGCAAGCGGTGA 840  
|  
|  
|  
QY 841 AGTTGACCGGCTGGGAGTAGACGATCGCAAGATTAAACTCAAGAGAAATGACGGGAGCC 900  
|  
|  
|  
Db 841 AGTTGACCGGCTGGGAGTAGACGATCGCAAGATTAAACTCAAGAGAAATGACGGGAGCC 900  
|  
|  
|  
QY 901 GCACAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 960  
|  
|  
|  
Db 901 GCACAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 960  
|  
|  
|  
QY 961 GACATGAGCGAATTTTCTAGAGATAGATTAGTGC---TTGCGGAAGCTTAAACAGAGT 1017  
|  
|  
|  
Db 961 GACATGAGCGAATTTTCTAGAGATAGATTAGTGC---TTGCGGAAGCTTAAACAGAGT 1017  
|  
|  
|  
QY 1018 CTGATAGGCTGTGCTACGCTCGTGTGATGATGTTAGTGGTTAGTCCGCAACGAGGCGCA 1077  
|  
|  
|  
Db 1018 CTGATAGGCTGTGCTACGCTCGTGTGATGATGTTAGTGGTTAGTCCGCAACGAGGCGCA 1077  
|  
|  
|  
QY 1021 CTGATAGGCTGTGCTACGCTCGTGTGATGATGTTAGTGGTTAGTCCGCAACGAGGCGCA 1080  
|  
|  
|  
Db 1021 CTGATAGGCTGTGCTACGCTCGTGTGATGATGTTAGTGGTTAGTCCGCAACGAGGCGCA 1080  
|  
|  
|  
QY 1081 ACCCTTGTCAATTAATGCTCAATCTTGTGGGCACTTTAATGAGACTGCGCGGTGACAA 1137  
|  
|  
|  
Db 1081 ACCCTTGTCAATTAATGCTCAATCTTGTGGGCACTTTAATGAGACTGCGCGGTGACAA 1137  
|  
|  
|  
QY 1138 CCGAGGAAAGTGGGAGTGAACGTCAAGTCTCATGCGCTTATAGGGAGCTTCAACG 1197  
|  
|  
|  
Db 1141 CCGAGGAAAGTGGGAGTGAACGTCAAGTCTCATGCGCTTATAGGGAGCTTCAACG 1200  
|  
|  
|  
QY 1198 TAATACAAATGGCCGCTTACAGAGGGTGTCCAAACCCGAGAGGGAGCTTAATCTCAGAAAGC 1257  
|  
|  
|



Db 1201 TAATACATGCGCGCTACAGAGGGTTGCCAACCCGGAGGGAGGCTAATCTCAGAAAGC 1260  
Qy 1258 GCGTGTAGTCCGGATCGAGAGTCTGCACTCGGTAAGTCCGAATCGCTAGTAAT 1317  
Db 1261 GGGTCTAGTCCGGATCGAGAGTCTGCACTCGGTAAGTCCGAATCGCTAGTAAT 1320  
Qy 1318 CCGGATCGAGAGTCCGGATCGAGAGTCTGCACTCGGTAAGTCCGAATCGCTAGTAAT 1377  
Db 1321 CCGGATCGAGAGTCCGGATCGAGAGTCTGCACTCGGTAAGTCCGAATCGCTAGTAAT 1380  
Qy 1378 CATGGAGTGGGTTTCCAGAGAGAGTGTCTAACTGTAAGAGAGGCGCTTCCACCG 1437  
Db 1381 CATGGAGTGGGTTTCCAGAGAGAGTGTGTAACTGTAAGAGAGGCGCTTCCACCG 1440  
Qy 1438 TGAGATTCACTAGTGGGTTG 1457  
Db 1441 TGAGATTCACTAGTGGGTTG 1460

RESULT 52  
AY123805 1497 bp DNA linear BCT 23-SBP-2003  
LOCUS Nitrosospira sp. Nv6 16S ribosomal RNA gene, partial sequence.  
DEFINITION AY123805  
ACCESSION AY123805  
VERSION AY123805.1 GI:24474434  
KEYWORDS Nitrosospira sp. Nv6  
SOURCE Nitrosospira sp. Nv6  
ORGANISM Nitrosospira sp. Nv6  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospira.  
1 (bases 1 to 1497)  
Purkhof, U., Wagner, M., Timmermann, G., Pommerening-Roeser, A. and  
Koop, H. P.  
16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonade  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
131310037  
2 (bases 1 to 1497)  
Purkhof, U., Timmermann, G., Koop, H. P., Pommerening-Roeser, A. and  
Wagner, M.  
Direct Submission  
Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
Location/Qualifiers  
1. 1497  
/organism="Nitrosospira sp. Nv6"  
/mol\_type="genomic DNA"  
/strain="Nv6"  
/db\_xref="taxon:136549"  
<1. >1497  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 88.4%; Score 1287.6; DB 1; Length 1497;  
Best Local Similarity 93.6%; Pred. No. 3.5e-13;  
Matches 1366; Conservative 0; Mismatches 89; Indels 5; Gaps 2;

Qy 1 ATTGAACGCTGGCGGATCTTTACATGCAAGTGAAACGCGACGAGATGCTTGCAT 60  
Db 3 ATTGAACGCTGGCGGATCTTTACATGCAAGTGAAACGCGACGAGATGCTTGCAT 60  
Qy 61 CTGGTGGCGAGTGGCGGAGCGGTGAGTAATGCATCGGAACTGTAAGAGAGAGGAGTA 120  
Db 61 CTGGTGGCGAGTGGCGGAGCGGTGAGTAATGCATCGGAACTGTAAGAGAGAGGAGTA 120  
Qy 121 ACGCATCGAAGATGCTAATACGCAATATCTTAAGAGAGAGAGAGAGAGAGAGTA 180  
Db 121 ACGCATCGAAGATGCTAATACGCAATATCTTAAGAGAGAGAGAGAGAGAGAGTA 180  
Qy 181 GACCTTGCCTTTTGGAGCGCGCATGTCTGATTAGCTAGTTGGTGGTAAAGGCTTAC 240  
Db 181 GACCTTGCCTTTTGGAGCGCGCATGTCTGATTAGCTAGTTGGTGGTAAAGGCTTAC 240

Qy 241 CAAGCGACGATCAGTATGTTGCTCTGAGAGAGACGACGCGCACTGGGACTGAGACAGC 300  
Db 241 CAAGCGATGATCAGTATGTTGCTCTGAGAGAGACGACGCGCACTGGGACTGAGACAGC 300  
Qy 301 GCCCAGACTCTTACCGGAGAGCAGCAGTGGGGAATTTTGGACAAATGAGCGCAAGCTGATC 360  
Db 301 GCCCAGACTCTTACCGGAGAGCAGCAGTGGGGAATTTTGGACAAATGAGCGCAAGCTGATC 360  
Qy 361 CAGCAATGCGCGTGAAGAGAGAGGCTTCCGGTTTGAAGCTCTTTCACTGAGAGAGA 420  
Db 361 CAGCAATGCGCGTGAAGAGAGAGGCTTCCGGTTTGAAGCTCTTTCACTGAGAGAGA 420  
Qy 421 AAAGTTACGGTAATTAATGTAACCAATGACCGTATGACAGAGAGAGCAGCGCTAAC 480  
Db 421 AAAGTTACGGTAATTAATGTAACCAATGACCGTATGACAGAGAGAGCAGCGCTAAC 480  
Qy 481 TACGTGCGACAGCGCGGTAAATACGTAGAGGTGCAAGGCTTAATCGAATTACTGGCGT 540  
Db 481 TACGTGCGACAGCGCGGTAAATACGTAGAGGTGCAAGGCTTAATCGAATTACTGGCGT 540  
Qy 541 AAAGGTGCGACAGCGCGGTAAATACGTAGAGGTGCAAGGCTTAATCGAATTACTGGCGT 600  
Db 541 AAAGGTGCGACAGCGCGGTAAATACGTAGAGGTGCAAGGCTTAATCGAATTACTGGCGT 600  
Qy 601 TGGCTTTGAAACTAAGAGCTAGAGTGGCAGAGAGAGTGAATTCATGTTAGAGAG 660  
Db 601 TGGCTTTGAAACTAAGAGCTAGAGTGGCAGAGAGAGTGAATTCATGTTAGAGAG 660  
Qy 661 TGAATGCGTGAAGATATGAAAGAACATGATGCGAAGAGCGCTCTGGGTTAACT 720  
Db 661 TGAATGCGTGAAGATATGAAAGAACATGATGCGAAGAGCGCTCTGGGTTAACT 720  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGGTACTCAAGCC 780  
Db 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGGTACTCAAGCC 780  
Qy 781 CTAAACGATGCACTAGTGTGGGCTTTAGGCTTGTGTAAGAGCTAAACGCGTGA 840  
Db 781 CTAAACGATGCACTAGTGTGGGCTTTAGGCTTGTGTAAGAGCTAAACGCGTGA 840  
Qy 841 AGTTACCGCTCGGGAGAGTACGCTGCAAGATTAACTCAAGAGAAATTGACGGGAGCC 900  
Db 841 AGTTACCGCTCGGGAGAGTACGCTGCAAGATTAACTCAAGAGAAATTGACGGGAGCC 900  
Qy 901 GCAAGAGCGGTGATTAATGTTGATTAATTCGATGCAAGCGGAAACCTTACCTTACCCTT 960  
Db 901 GCAAGAGCGGTGATTAATGTTGATTAATTCGATGCAAGCGGAAACCTTACCTTACCCTT 960  
Qy 961 GACATGTAGCGAATTTTCTAGAGATAGATTGTC---TTGGGGAACGCTAACACAGGTTG 1017  
Db 961 GACATGTAGCGAATTTTCTAGAGATAGATTGTC---TTGGGGAACGCTAACACAGGTTG 1020  
Qy 1018 CTGCATGCTGCTGCTCAGCTGCTGCTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1077  
Db 1021 CTGCATGCTGCTGCTCAGCTGCTGCTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1080  
Qy 1078 ACCCTTGTCAATTAATTGCCATCATTTGGTGGGCACTTTAATGAGATGCGGTGACAA 1137  
Db 1081 ACCCTTGTCAATTAATTGCCATCATTTGGTGGGCACTTTAATGAGATGCGGTGACAA 1140  
Qy 1138 CCGGAGGAAGGTGGGAGTGAAGTCAAGTCTCATAGGCGCTTAATGGGTAGGGCTTACACG 1197  
Db 1141 CCGGAGGAAGGTGGGAGTGAAGTCAAGTCTCATAGGCGCTTAAATGGGTAGGGCTTACACG 1200  
Qy 1198 TAATCAATGCGCTACAGAGGTTGCAACCCCGAGGGGAGCTAATCTCAGAAAGC 1257  
Db 1201 TAATCAATGCGCTACAGAGGTTGCAACCCCGAGGGGAGCTAATCTCAGAAAGC 1260  
Qy 1258 GCGTGTAGTCCGATCGAAGTCTGCACTGCACTCGGTGAAGTCCGAATGCTAGTAAT 1317  
Db 1261 GCGTGTAGTCCGATCGAAGTCTGCACTGCACTCGGTGAAGTCCGAATGCTAGTAAT 1320





RESULT 54  
NTL16SR  
LOCUS  
DEFINITION  
Nitrosolobus multiformis (ATCC 25196) 16S ribosomal RNA (16S rRNA)  
gene.  
ACCESSION  
135509  
135509.1 GI:530900  
KEYWORDS  
16S ribosomal RNA; ribosomal RNA; small subunit.  
SOURCE  
ORGANISM  
Nitrosospiira multiformis  
Nitrosospiira multiformis  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospiira.  
REFERENCE  
1 (bases 1 to 1478)  
Tejeda, A., Alm, E., Regan, J.M., Toze, S., Rittmann, B.E. and Stahl, D.A.  
Evolutionary relationships among ammonia- and nitrite-oxidizing  
bacteria  
J. Bacteriol. 176 (21), 6623-6630 (1994)  
JOURNAL  
PUBMED  
7961414  
COMMENT  
Original source text: Nitrosolobus multiformis (library: ATCC  
25196) DNA.  
FEATURES  
source  
Location/Qualifiers  
1..1478  
/organism="Nitrosospiira multiformis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1231"  
/rbsuse\_1ib="ATCC 25196"  
1..1478  
/gene="16S rRNA"  
1..1478  
/gene="16S rRNA"  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 88.2%; Score 1284.6; DB 1; Length 1478;  
Best Local Similarity 93.4%; Pred. No. 3.9e-13;  
Matches 1363; Conservative 0; Mismatches 92; Indels 5; Gaps 2;  
QY 1 ATTGAACGCTGGGCGCATGCTTTACACATGCAAGTGAACGCGACGACCGGATCTTGCAAT 60  
DB 21 ATTGAACGCTGGGCGCATGCTTTACACATGCAAGTGAACGCGACGCGGCGCAACCC-- 78  
QY 61 CTGGTGGCGATGGCGGACGCGTGAATGATCGGAACGTAATCCAGAGAGGCGGAT 120  
DB 79 CTGGTGGCGATGGCGGACGCGTGAATGATCGGAACGTAATCCAGAGAGGCGGAT 138  
QY 121 ACGCATGAAAGATGTGCTAATACCGCATATCTTAAAGAGGAAAGACAGGAGATGAAA 180  
DB 139 ACGCATGAAAGATGTGCTAATACCGCATATCTTAAAGAGGAAAGACAGGAGATGAAA 198  
QY 181 GACCTTGGCTTTTGAAGCGCGCATGCTGATTAAGTATGCTGAGGCGTAAAGCGCTAC 240  
DB 199 GACCTTGGCTTTTGAAGCGCGCATGCTGATTAAGTATGCTGAGGCGTAAAGCGCTAC 258  
QY 241 CAAGGCGATGATGATGATGCTGAGAGAGACGACGACCACTGGAGCTGAGACAG 300  
DB 259 CAAGGCGATGATGATGATGCTGAGAGAGACGACGACCACTGGAGCTGAGACAG 318  
QY 301 GCCCAAGCTCTTACGGGAGGACGACGATGGGAAATTTTGAACAATGGGCGCAAGCTGATC 360  
DB 319 GCCCAAGCTCTTACGGGAGGACGACGATGGGAAATTTTGAACAATGGGCGCAAGCTGATC 378  
QY 361 CAGCAATGCGCGCTGATGAGAAAGCGCTTCCGGGTTGTAAGCTTTTCACTGAGAAAC 420  
DB 379 CAGCAATGCGCGCTGATGAGAAAGCGCTTCCGGGTTGTAAGCTTTTCACTGAGAAAC 438  
QY 421 AAAGTTACGCTTAATATGATGACCATGACCGTATGACAGAAAGAGACCGGCTAAC 480  
DB 439 AAAGTTACGCTTAATATGATGACCATGACCGTATGACAGAAAGAGACCGGCTAAC 498  
QY 481 TACGTCCAGACGCGCGGATTAATGATGAGGCTCAAGCGTTAATCGAATTAATCTGAGCG 540  
DB 499 TACGTCCAGACGCGCGGATTAATGATGAGGCTCAAGCGTTAATCGAATTAATCTGAGCG 558

QY 541 AAAGGTGCGACGCGCGCTTGTAAATGATGATGTAATCCCGGCTTAACCTGGGAAT 600  
DB 559 AAAGGTGCGACGCGCGCTTGTAAATGATGATGTAATCCCGGCTTAACCTGGGAAT 618  
QY 601 TGGCTTGAATCAACAAGCTAGATGAGGACGAGAGGATGGAATTCATGTGACAG 660  
DB 619 TGGCTTGAATCAACAAGCTAGATGAGGACGAGAGGATGGAATTCATGTGACAG 678  
QY 661 TGAATATGCTAGATGATGAGAAACATGATGAGGACGAGGATGGAATTCATGTGAC 720  
DB 679 TGAATATGCTAGATGATGAGAAACATGATGAGGACGAGGATGGAATTCATGTGAC 738  
QY 721 GACGCTATGACGAAAGCGTGGGAGCAACAGATTAAGTATACCTGATGACAGCC 780  
DB 739 GACGCTATGACGAAAGCGTGGGAGCAACAGATTAAGTATACCTGATGACAGCC 798  
QY 781 CTAAACGATGCACTAGTATGTTGGGCTTATTAAGCTTGTGTAACGAACTAACCGGTGA 840  
DB 799 CTAAACGATGCACTAGTATGTTGGGCTTATTAAGCTTGTGTAACGAACTAACCGGTGA 858  
QY 841 AGTTGACCGCTGGGAGATACGCTGCAAGATTAATCAAAAGAAATGACGGGAGACC 900  
DB 859 AGTTGACCGCTGGGAGATACGCTGCAAGATTAATCAAAAGAAATGACGGGAGACC 918  
QY 901 GCACAAAGCGGATGATGATGATTAATTCAGATGCAACGCAAAACCTTAACCTACCT 960  
DB 919 GCACAAAGCGGATGATGATGATTAATTCAGATGCAACGCAAAACCTTAACCTACCT 978  
QY 961 GACATGTACGAAATTTTCTAGAGATGATTAAGTATGCT--TTGGGAAACGCTAACAGGTG 1017  
DB 979 GACATGTACGAAATTTTCTAGAGATGATTAAGTATGCT--TTGGGAAACGCTAACAGGTG 1038  
QY 1018 CTGCATGCTGTCGTCACGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1077  
DB 1039 CTGCATGCTGTCGTCACGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1098  
QY 1078 ACCCTTGTATTAATTTGATCATATTTGATGATGATGATGATGATGATGATGATGATG 1137  
DB 1099 ACCCTTGTATTAATTTGATCATATTTGATGATGATGATGATGATGATGATGATGATG 1158  
QY 1138 CCGGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1197  
DB 1159 CCGGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1218  
QY 1198 TAATCAATGAGCGCTGACAGAGGTTGCCAACCCGAGGAGGAGCTATTCAGAAAGC 1257  
DB 1219 TAATCAATGAGCGCTGACAGAGGTTGCCAACCCGAGGAGGAGCTATTCAGAAAGC 1278  
QY 1258 GGGTGTAGTCCGATGCGAGTCTGCAACTCGACTCCGTAAGTGGAAATGCTAGTAAT 1317  
DB 1279 GGGTGTAGTCCGATGCGAGTCTGCAACTCGACTCCGTAAGTGGAAATGCTAGTAAT 1338  
QY 1318 CCGGATGACATGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377  
DB 1339 CCGGATGACATGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1398  
QY 1378 CATGGAGTGGGTTTCAACAAGACGATGATGATGATGATGATGATGATGATGATGATG 1437  
DB 1399 CATGGAGTGGGTTTCAACAAGACGATGATGATGATGATGATGATGATGATGATGATG 1458  
QY 1438 TGAGATTCATGACTGGGGTG 1457  
DB 1459 TGAGATTCATGACTGGGGTG 1478

RESULT 55  
AF353163  
LOCUS  
DEFINITION  
Nitrosospiira sp. PJA1 16S ribosomal RNA gene, partial sequence.  
ACCESSION  
AF353163  
AF353163.1 GI:13487772  
KEYWORDS  
SOURCE  
Nitrosospiira sp. PJA1

ORGANISM	Nitrosospora sp. PJA1
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosospora.
AUTHORS	1 (bases 1 to 1457)
TITLE	Takahashi, R., Sato, K. and Tokuyama, T.
JOURNAL	Nitrosolobus sp. PJA1 16S rRNA gene
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 1457)
TITLE	Takahashi, R., Sato, K. and Tokuyama, T.
JOURNAL	Direct Submission
FEATURES	Submitted (26-FEB-2001) Agricultural and Biological Chemistry, College of Bioresource Sciences, Nihon University, 1866 Kameino, Fujisawa 252-8510, Japan
Source	Location/Qualifiers
	1..1457
	/organism="Nitrosospora sp. PJA1"
	/mol_type="genomic DNA"
	/db_xref="taxon:155921"
	/note="similar to Nitrosospora (Nitrosolobus) multiformis"
	similar to Nitrosospora (Nitrosolobus) multiformis"
	<1..>1457
	/product="16S ribosomal RNA"
ORIGIN	
Query Match	88.2%; Score 1284.4; DB 1; Length 1457;
Best Local Similarity	93.4%; Pred. No. 4e-13; Indels 4; Gaps 2;
Matches 1363; Conservative	0; Mismatches 92;
QY	1 ATTTGAACCTGGCGCGCATGCTTTTACACATGCAAGTCGAACCGCAGACGAGTCTTGAT 60
DB	1 ATTTGAACCTGGCGCGCATGCTTTTACACATGCAAGTCGAACCGCAGACGAGGCGCAAC-- 58
QY	61 CTGTGGCGAGTGGCGGACGGGTGATGTAATGCATGGAACGTAATCCAGAAAGAGGGGGGTA 120
DB	59 CTGTGGCGAGTGGCGGAAACGGGTGATGTAATGCATGGAACGTAATCCAGAAAGAGGGGTA 118
QY	121 AGCGATCCGAAATGTCCTAATACCGCATATCTTAAAGAGAAAGACGGGGATTCGAAA 180
DB	119 AGCGATCCGAAAGTGTGCTAATACCGCATATCTTAAAGAGAAAGACGGGGATTCGAAA 178
QY	181 GACCTTGGCGCTTTGGAGCGCGCCAGTGTCTGATTAGCTAGTTGGTGGGTAAAGGCTTAC 240
DB	179 GACCTTGGCGCTTTGGAGCGCGCCAGTGTCTGATTAGCTAGTTGGTGGGTAAAGGCTTAC 238
QY	241 CAAGGCGAGCATCAGTAGTTGTTCTGAGAGAGACGACACCCACACTGGGACTGAGACACG 300
DB	239 CAAGGCGTGCATCAGTAGTTGTTCTGAGAGAGACGACACCCACACTGGGACTGAGACACG 298
QY	301 GCCCAGACTCTTACCGGAGGACGACATGGGGAAATTTTGGACAAATGGGCGCAGACCTGATC 360
DB	299 GCCCAGACTCTTACCGGAGGACGACATGGGGAAATTTTGGACAAATGGGCGCAGACCTGATC 358
QY	361 CAGCAATGGCGGTGATGTAAGAAAGGCTTCGGGTGTAAGGCTCTTTGAGTCGAGAGAGA 420
DB	359 CAGCAATGGCGGTGATGTAAGAAAGGCTTCGGGTGTAAGGCTCTTTGAGTCGAGAGAGA 418
QY	421 AAAGTTACGTTAAT 480
DB	419 AAAGTTACGTTAAT 478
QY	481 TACGTCCAGACAGCCCGGTTAATAGTAGGTTCAGACGTTAATTCGAAATTTACGGGCGCT 540
DB	479 TACGTCCAGACAGCCCGGTTAATAGTAGGTTCAGACGTTAATTCGAAATTTACGGGCGCT 538
QY	541 AAAGGGTGGCAGGCGGCTTTTGAATGTCAGATGTGAATATCCCGGGCTTAACCTGGGAAT 600
DB	539 AAAGGGTGGCAGGCGGCTTTTGAATGTCAGATGTGAATATCCCGGGCTTAACCTGGGAAT 598
QY	601 TGGCTTTGAATCAAAAGCTAGAGTGTGGCAGAGGAGAGGTGGAATTTCAATGTGAGCAG 660
DB	599 TGGCTTTGAATCAAAAGCTAGAGTGTGGCAGAGGAGAGGTGGAATTTCAACGTGTGAGCAG 658
QY	661 TGAATATGCTAGAGATATGGAAGAACATGCATGGCCAGAGCAGCTTCCTGGTTAACTACT 720

Db	659	TGAAATGCGTAGAGATGTGAGGAAACCCGATGGCGAAGCAGCCCTCGGGTTAACT	718
QY	721	GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTAGTCCAGCC	780
Db	719	GACGCTCAGGCAAGAAAGGTGGGGAGCAAAACAGATTAGATACCTGTAGTCCAGCC	778
QY	781	CTAAACGATGTCAACTGATTGTTGGGCTTTATTAGGCTTGTATACGAACACTAAACGGGTGA	840
Db	779	CTAAACGATGTCAACTGATTGTTGGGCTTTATTAGGCTTGTATACGAACACTAAACGGGTGA	838
QY	841	AGTTGACCCGCTGGGGAGTACCGGTCCGAAGATTAAACTCAAGAAATTGACGGGGACCC	900
Db	839	AGTTGACCCGCTGGGGAGTACCGGTCCGAAGATTAAACTCAAGAAATTGACGGGGACCC	898
QY	901	GCAACAAGCGGTGATTATGTGATTAAATTCGATGCAACGGGAAAACTTTACTTACCCTT	960
Db	899	GCAACAAGCGGTGATTATGTGATTAAATTCGATGCAACGGGAAAACTTTACTTACCCTT	958
QY	961	GACATGT--AGCGAATTTTCTAGAGATAGTTAGTGCTTCGGGAAAGCTTACACAGGTGC	1018
Db	959	GACATGTGCCGAGACCCGCTGAANAAGTGGGCGTCCCGAAGAAACGGCAACAGGTGC	1018
QY	1019	TGCATGCTGTGCGTACAGCTGCTGTCGAGATGTTGGGTTAAAGTCCGCAACGAGCGCAA	1078
Db	1019	TGCATGCTGTGCGTACAGCTGCTGTCGAGATGTTGGGTTAAAGTCCGCAACGAGCGCAA	1078
QY	1079	CCCTTGTCAATTAAATTCGATCAATTTGGTTGGGCACTTTAATGAGACTGCCGTGACAAAC	1138
Db	1079	CCCTTGTCAATTAAATTCGATCAATTTGGTTGGGCACTTTAATGAGACTGCCGTGACAAAC	1138
QY	1139	CGGAGAAAGTGGGGAATGACGTCAAGTCCCTCATGCGCCCTTAATGGGTAGGGCTTCAACGT	1198
Db	1139	CGGAGAAAGTGGGGAATGACGTCAAGTCCCTCATGCGCCCTTAATGGGTAGGGCTTCAACGT	1198
QY	1199	AATACAAATGGGCGGCTACAGAGGGTGGCAACCCGCGAGGGGAGACTAATCTCAAGAAACG	1258
Db	1199	AATACAAATGGGCGGCTACAGAGGGTGGCAACCCGCGAGGGGAGACTAATCTCAAGAAACG	1258
QY	1259	CGTGTAGTCCGATCGGAATCGGAATCTGCAACTCGTGAAGTCCGAAATGCTTAGTAAATC	1318
Db	1259	CGTGTAGTCCGATCGGAATCGGAATCTGCAACTCGTGAAGTCCGAAATGCTTAGTAAATC	1318
QY	1319	GCGGATCAAGCATATGTCGGGTGAATACGTTCCCGGCTTTGTACACACGCCCGTACACC	1378
Db	1319	GCGGATCAAGCATATGTCGGGTGAATACGTTCCCGGCTTTGTACACACGCCCGTACACC	1378
QY	1379	ATGGAGATGGGTTTACACAGAACGAGTAGTCTTAACCGTAAGAGAGGCGCTTGCAACGT	1438
Db	1379	ATGGAGATGGGTTTACACAGAACGAGTAGTCTTAACCGTAAGAGAGGCGCTTGCAACGT	1438
QY	1439	GAGATTCAATGACTGGGGTG	1457
Db	1439	GAGATTCAATGACTGGGGTG	1457
RESULT 56			
AF386756		1486 bp	DNA linear ENV 10-MAY-2004
LOCUS	AF386756		
DEFINITION	Nitrosospiira sp. R3c5	16S ribosomal RNA gene, partial sequence.	
ACCESSION	AF386756		
VERSION	AF386756.1	GI:17864833	
KEYWORDS	ENV.		
SOURCE	Nitrosospiira sp. R3c5		
ORGANISM	Nitrosospiira sp. R3c5		
REFERENCE	Burrell, P. C., Phalen, C. M. and Hovanec, T. A.		
AUTHORS	Identification of Bacteria responsible for ammonia oxidation in		
TITLE	freshwater aquaria		
JOURNAL	Appl. Environ. Microbiol.	67 (12), 5791-5800 (2001)	
PUBMED	11722936		

REFERENCE 2 (bases 1 to 1486)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria Group, 6100 Condon Dr., Moorpark, CA 93021, USA  
FEATURES  
source  
1. 1486  
/organism="Nitrosospiria sp. R3c5"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated from freshwater aquaria enrichment"  
/db\_xref="taxon:180550"  
/clone="R3c5"  
/environmental\_sample  
<1..>1486  
/product="16S ribosomal RNA"  
ORIGIN  
RNA  
Query Match 88.2%; Score 1284.4; DB 3; Length 1486;  
Best Local Similarity 93.4%; Pred. No. 4e-13;  
Matches 1364; Conservative 0; Mismatches 91; Indels 5; Gaps 2;  
QY 1 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCAACGCGACGACGCGATGCTTGCAAT 60  
DB 15 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCAACGCGACGACGCGGCGCAAC-- 72  
QY 61 CTGGTGGCGAGTGGCGGACGCGGTGATTAATGCAATCGGAACTGATCCAGAAAGGGGGGTA 120  
DB 73 CTGGTGGCGAGTGGCGGACGCGGTGATTAATGCAATCGGAACTGATCCAGAAAGGGGGGTA 132  
QY 121 AGCATCGAAAGATGTGCTAATACCGCATTAATCTTAAGGAGAAAGACGAGGATCGAAA 180  
DB 133 AGCATCGAAAGATGTGCTAATACCGCATTAATCTTAAGGAGAAAGACGAGGATCGCAA 192  
QY 181 GACCTTGCGCTTTGGAGCGCGCGATGTCTGATTAGCTAGTGGGTAAAGGCTTAC 240  
DB 193 GACCTTGCGCTTTGGAGCGCGCGATGTCTGATTAGCTAGTGGGTAAAGGCTTAC 252  
QY 241 CAAGGGAAGATGATGATGTTGCTGAGAGAGACACACGACCTGGGATCGAGACAG 300  
DB 253 CAAGGGAAGATGATGATGTTGCTGAGAGAGACACACGACCTGGGATCGAGACAG 312  
QY 301 GCCCAACTCTTACGGGAGGACAGTGGGGAATTTTGGCAATGGGCGCAAGCTGATC 360  
DB 313 GCCCAACTCTTACGGGAGGACAGTGGGGAATTTTGGCAATGGGCGCAAGCTGATC 372  
QY 361 CAGCAATGCGCGCTGAGTGAAGAGGCTTGGGTTGTAAGCTTTTCAAGTCGAGAGA 420  
DB 373 CAGCAATGCGCGCTGAGTGAAGAGGCTTGGGTTGTAAGCTTTTCAAGTCGAGAGA 432  
QY 421 AAAGTTAAGGTAAATATCTGTAACCATGACCGGTATCCAGAAAGACACCGGCTAAC 480  
DB 433 AAAGTTAAGGTAAATATCTGTAACCATGACCGGTATCCAGAAAGACACCGGCTAAC 492  
QY 481 TAGGTGCGACAGCCCGGTAATACGTAAGGTGCAAGCGTTAATCGAATTAATCTGAGCGT 540  
DB 493 TAGGTGCGACAGCCCGGTAATACGTAAGGTGCAAGCGTTAATCGAATTAATCTGAGCGT 552  
QY 541 AAAGGTGCGACAGCCCGCTTGTAACTGATGTAATATCCCGGCTTAACTTGGGAAT 600  
DB 553 AAAGGTGCGACAGCCCGCTTGTAACTGATGTAATATCCCGGCTTAACTTGGGAAT 612  
QY 601 TGGCTTTGAAACTAACAAGCTAGATGTGCGACAGGAGTGAATTTCCATGTGTAGCAG 660  
DB 613 TGGCTTTGAAACTAACAAGCTAGATGTGCGACAGGAGTGAATTTCCATGTGTAGCAG 672  
QY 661 TGAATATGCTAAGATATGAAAGACATGATGCGAGGACGCTCTGAGTTAACTACT 720  
DB 673 TGAATATGCTAAGATATGAAAGACATGATGCGAGGACGCTCTGAGTTAACTACT 732  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTTGGTATGTCACGCG 780  
DB 733 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTTGGTATGTCACGCG 792

QY 781 CTTAAAGATGCACTAGTGTGGGCTTATATAGGCTTGTAAGCAAGCTAACGGGTGA 840  
DB 793 CTTAAAGATGCACTAGTGTGGGCTTATATAGGCTTGTAAGCAAGCTAACGGGTGA 852  
QY 841 AGTTGACCGCTGGGAGATGAGGTGCAAGATTAATACTCAAGAAATTTGACGGGAGCC 900  
DB 853 AGTTGACCGCTGGGAGATGAGGTGCAAGATTAATACTCAAGAAATTTGACGGGAGCC 912  
QY 901 GCAACAAGCGGTGATTAATGATGATTAATTCATGCAAGCGGAAAACTTACTTACCCTT 960  
DB 913 GCAACAAGCGGTGATTAATGATGATTAATTCATGCAAGCGGAAAACTTACTTACCCTT 972  
QY 961 GACATGATGCAATTTTCTAAGATATGATTAAGTGC---TTGGGAAAGCTTAACAGAGTG 1017  
DB 973 GACATGATGCAAGTCCCGGAGAGGTGGGTGTGCCCCAAAGGAGCGGTAACAGAGTG 1032  
QY 1018 CTGCATGCTGTGCTGACGCTGCTGTGCTGATGATGTTGGTTAAGTCCCGCAACGAGCGCA 1077  
DB 1033 CTGCATGCTGTGCTGACGCTGCTGTGCTGATGATGTTGGTTAAGTCCCGCAACGAGCGCA 1092  
QY 1078 ACCCTGTGCTAATTTGCTGATATTTGGTGGGCACTTTAATGACACTGCCGTGACAAA 1137  
DB 1093 ACCCTGTGCTAATTTGCTGATATTTGGTGGGCACTTTAATGACACTGCCGTGACAAA 1152  
QY 1138 CCGGAGAAAGGTGGGATGACGTCAGATCTCATGAGCCCTTAATGGGTAGGGCTTACACAG 1197  
DB 1153 CCGGAGAAAGGTGGGATGACGTCAGATCTCATGAGCCCTTAATGGGTAGGGCTTACACAG 1212  
QY 1198 TAAATCAATGCGCTACAGAGGTTGGCCAAACCGCGAGGGGAGCTAATCTCAGAAAGC 1257  
DB 1213 TAAATCAATGCGCTACAGAGGTTGGCCAAACCGCGAGGGGAGCTAATCTCAGAAAGC 1272  
QY 1258 GCGTGTGATCCGATTCGAGATCTGCACTGCACTCCGTGAAGTGGAAATGCTTATAT 1317  
DB 1273 GCGTGTGATCCGATTCGAGATCTGCACTGCACTCCGTGAAGTGGAAATGCTTATAT 1332  
QY 1318 GCGGATCAGATGATCGGATGATACGTTCCGGGCTTGTGACACACCGCCGTACAC 1377  
DB 1333 GCGGATCAGATGATCGGATGATACGTTCCGGGCTTGTGACACACCGCCGTACAC 1392  
QY 1378 CATGGAGTGGGTTTCAACAGAGAGGTAGTCTAACCGTAAAGAGGGCGCTTGCACAG 1437  
DB 1393 CATGGAGTGGGTTTCAACAGAGAGGTAGTCTAACCGTAAAGAGGGCGCTTGCACAG 1452  
QY 1438 TGAGATTCAATGACTGGGCTG 1457  
DB 1453 TGAGATTCAATGACTGGGCTG 1472  
RESULT 57  
AY123788 1497 bp DNA linear BCT 23-SEP-2003  
LOCUS Nitrosospiria sp. Nsp41 16S ribosomal RNA gene, partial sequence.  
DEFINITION  
ACCESSION AY123788  
VERSION AY123788.1 GI:24474417  
KEYWORDS  
SOURCE  
ORGANISM Nitrosospiria sp. Nsp41  
Nitrosospiria sp. Nsp41  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospiria.  
REFERENCE  
AUTHORS Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Roser, A. and  
Koops, H.-P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
JOURNAL Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
PUBMED 13130037  
REFERENCE 2 (bases 1 to 1497)  
AUTHORS Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Roser, A. and  
Wagner, M.  
TITLE Direct Submission

JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany

FEATURES  
Location/Qualifiers

1..1497  
/organism="Nitrosospirila sp. Nsp17"  
/mol\_type="genomic DNA"  
/strain="Nsp41"  
/db\_xref="taxon:200115"  
<1..>1497  
/product="16S ribosomal RNA"

## ORIGIN

Query Match 88.2%; Score 1284.4; DB 1; Length 1497;

Best Local Similarity 93.4%; Pred. No. 3.9e-13;

Matches 1364; Conservative 0; Mismatches 91; Indels 5; Gaps 2;

1 ATTGAACGCTGGCGGCGATGCTTTACATGCAAGTCGAAACGGACGACGAGATGCTTGAT 60  
3 ATTGAACGCTGGCGGCGATGCTTTACATGCAAGTCGAAACGGACGACGAGATGCTTGAT 60  
61 CTGGTGCGAGTGGCGGCGAGCGGTGATGATGATCGGAAAGTATCCAGAAAGCGGCGGATA 120  
61 CTGGTGCGAGTGGCGGCGAGCGGTGATGATGATCGGAAAGTATCCAGAAAGCGGCGGATA 120  
121 ACCGATCGAAAGATGCTGCTAATACCGCATATCTCTAAGAGAAAGCAAGCGGATCGAAA 180  
121 ACCGATCGAAAGATGCTGCTAATACCGCATATCTCTAAGAGAAAGCAAGCGGATCGAAA 180  
181 GACCTTCGCTTTTGAAGCGCGCGATGCTGATTAAGTATGTTGGGTGATTAAGCGCTTAC 240  
181 GACCTTCGCTTTTGAAGCGCGCGATGCTGATTAAGTATGTTGGGTGATTAAGCGCTTAC 240  
181 GACCTTCGCTTTTGAAGCGCGCGATGCTGATTAAGTATGTTGGGTGATTAAGCGCTTAC 240  
241 CAAGCGACGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
241 CAAGCGACGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
241 CAAGCGACGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
301 GCCGAGACTCTTAACGCGGAGCGAGCGATGCGGAAATTTTGAACAATGGGCGAAACCTGATC 360  
301 GCCGAGACTCTTAACGCGGAGCGAGCGATGCGGAAATTTTGAACAATGGGCGAAACCTGATC 360  
301 GCCGAGACTCTTAACGCGGAGCGAGCGATGCGGAAATTTTGAACAATGGGCGAAACCTGATC 360  
361 CAGCATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
361 CAGCATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
361 CAGCATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
421 AAAGGTTACGCTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
421 AAAGGTTACGCTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
421 AAAGGTTACGCTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
481 TACGTGCGAGCGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
481 TACGTGCGAGCGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
481 TACGTGCGAGCGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
541 AAAGGTTACGCTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
541 AAAGGTTACGCTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
541 AAAGGTTACGCTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
601 TCGGTTTGAATCTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
601 TCGGTTTGAATCTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
601 TCGGTTTGAATCTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
661 TGAATGCGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
661 TGAATGCGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
661 TGAATGCGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
721 GACGCTCATGACGAAAGCGTGGGAGCAACAAGATTAATGATCCCTGGTATCCAGCC 780  
721 GACGCTCATGACGAAAGCGTGGGAGCAACAAGATTAATGATCCCTGGTATCCAGCC 780  
721 GACGCTCATGACGAAAGCGTGGGAGCAACAAGATTAATGATCCCTGGTATCCAGCC 780  
781 CTAAACGATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
781 CTAAACGATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
841 AGTTACCGGCTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900

841 AGTTACCGGCTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
901 GCACAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
901 GCACAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
961 GACATGTACGAAAGCGGCGGAGAGTGGGTGATGATGATGATGATGATGATGATGATGATGAT 1017  
961 GACATGTACGAAAGCGGCGGAGAGTGGGTGATGATGATGATGATGATGATGATGATGATGAT 1017  
961 GACATGTACGAAAGCGGCGGAGAGTGGGTGATGATGATGATGATGATGATGATGATGATGAT 1020  
1018 CTGCATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077  
1021 CTGCATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
1078 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
1081 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
1138 CCGAGGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
1141 CCGAGGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
1198 TAATACAAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257  
1201 TAATACAAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
1258 GCGTGTGATGCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317  
1261 GCGTGTGATGCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1318 CCGGAT 1377  
1321 CCGGAT 1380  
1378 CATGGAGTGGGTTTACCGAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1437  
1381 CATGGAGTGGGTTTACCGAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
1438 TGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
1441 TGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1460

RESULT 58  
AY123804 1497 bp DNA linear BCT 23-SBP-2003  
LOCUS  
DEFINITION Nitrosospirila sp. Nsp17 16S ribosomal RNA gene, partial sequence.  
ACCESSION AY123804  
VERSION AY123804.1 GI:24474433  
KEYWORDS  
SOURCE  
ORGANISM Nitrosospirila sp. Nsp17  
Nitrosospirila sp. Nsp17  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospirila.  
REFERENCE  
AUTHORS 1 (bases 1 to 1497)  
Purkhold, U., Wagner, M., Timmermann, G., Pommerening-Rosser, A. and  
Koops, H. P.  
TITLE 16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial  
ammonia-oxidizing isolates: extension of the dataset and proposal  
of a new lineage within the nitrosomonads  
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
REFERENCE  
AUTHORS Purkhold, U., Timmermann, G., Koops, H.-P., Pommerening-Rosser, A. and  
Wagner, M.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
FEATURES  
source  
1..1497  
/organism="Nitrosospirila sp. Nsp17"  
/mol\_type="genomic DNA"

/sra:in="Nsp17"  
/db\_xref="taxon:136546"  
<1..>1497  
/product="16S ribosomal RNA"

## ORIGIN

```
Query Match      88.2%; Score 1284.4; DB 1; Length 1497;
Best Local Similarity 93.4%; Pred. No. 3.9e-13;
Matches 1364; Conservative 0; Mismatches 91; Indels 5; Gaps 2;

QY      1 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCGAAACGCGACGAGCTTGGCAT 60
DB      3 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCGAAACGCGACGAGCTTGGCAT 60
QY      61 CTGGTGGCGAGTGGCGGACGGGTGATGATCGGAACGTATCCAGAAAGAGGGGGGTA 120
DB      61 CTGGTGGCGAGTGGCGGACGGGTGATGATCGGAACGTATCCAGAAAGAGGGGGGTA 120
QY      121 AGCGATCGAAAGATGCTAAATACCGCATATACCTTAAGAGGAAAGCAGGGGATCGAAA 180
DB      121 AGCGATCGAAAGATGCTAAATACCGCATATATCTTAAGAGGAAAGCAGGGGATCGCAA 180
QY      181 GACCTTGGCGCTTTTGGAGCGCGCATGTCTGATTAAGCTAGTGGTGAAGGCTTAC 240
DB      181 GACCTTGGCGCTTTTGGAGCGCGCATGTCTGATTAAGCTAGTGGTGAAGGCTTAC 240
QY      241 CAAGGCGAGATCAATGATGTTGTCTGAGAGAGACGACGACCACTGGGACTGAGACAG 300
DB      241 CAAGGCGATGATGATGATGTTGTCTGAGAGAGACGACGACCACTGGGACTGAGACAG 300
QY      301 GCCCACTCTCTAGGAGGAGCAGACAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATC 360
DB      301 GCCCACTCTCTAGGAGGAGCAGACAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATC 360
QY      361 CAGCAATGCGCGCTGATGAGAAAGGCTTGGGCTTGAAGCTTTTCACTGAGAGAG 420
DB      361 CAGCAATGCGCGCTGATGAGAAAGGCTTGGGCTTGAAGCTTTTCACTGAGAGAG 420
QY      421 AAAGGTTACGGTAAATATATGATGATCCATGACCGTATCCAGAAAGACACCGGCTTAC 480
DB      421 AAAGGTTACGGTAAATATATGATGATCCATGACCGTATCCAGAAAGACACCGGCTTAC 480
QY      481 TACGTCACAGACCGCGGATATACGTAAGGCTGCAAGCGTTAATCGGAATTAATCTGGAGCT 540
DB      481 TACGTCACAGACCGCGGATATACGTAAGGCTGCAAGCGTTAATCGGAATTAATCTGGAGCT 540
QY      541 AAAGGTCGCGACGCGCTTGTAGTCAATGATGTAATCCCGGCTTAACTTGGGAAT 600
DB      541 AAAGGTCGCGACGCGCTTGTAGTCAATGATGTAATCCCGGCTTAACTTGGGAAT 600
QY      601 TGGGTTTGAATCTACAAAGCTAGATGTGTGCAAGGAGGTGGAATTTCCATGTGTAGCAG 660
DB      601 TGGGTTTGAATCTACAAAGCTAGATGTGTGCAAGGAGGTGGAATTTCCATGTGTAGCAG 660
QY      661 TGAATATGCGTATGATATGGAAGAACATGATGCGCAAGGACGCTCTGGGTTAACTACT 720
DB      661 TGAATATGCGTATGATATGGAAGAACATGATGCGCAAGGACGCTCTGGGTTAACTACT 720
QY      721 GACGCTCATGACGAAGACGTGGGAGCAAAACAGATTAAGATACCTGGTATGTCACGCC 780
DB      721 GACGCTCATGACGAAGACGTGGGAGCAAAACAGATTAAGATACCTGGTATGTCACGCC 780
QY      781 CTAAACGATGTCATGATGTTGTGGCTTATTAAGCTTGTATACGAAGTAACTGGGTGA 840
DB      781 CTAAACGATGTCATGATGTTGTGGCTTATTAAGCTTGTATACGAAGTAACTGGGTGA 840
QY      841 AGTTGACGCGCTGGGAGTACGCTGCAAGATTAATACTCAAGAAATTGACGGGAGACC 900
DB      841 AGTTGACGCGCTGGGAGTACGCTGCAAGATTAATACTCAAGAAATTGACGGGAGACC 900
QY      901 GCAACAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGGAAAACTTAACTTACCTTT 960
DB      901 GCAACAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGGAAAACTTAACTTACCTTT 960
```

```
QY      961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC---TTGGGAAACGCTAACACAGGTG 1017
DB      961 GACATGTAGCGAAGCCCGCTGAGAGTGGGTGTGCCGAAGGACGGTTAACACAGGTG 1020
QY      1018 CTGCATGCTGTCTGTCAAGCTGTGTCTGTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1077
DB      1021 CTGCATGCTGTCTGTCAAGCTGTGTCTGTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1080
QY      1078 ACCCTTGTATTAATTGTCATCACTTTTGGTGGGCACTTTAAAGACATGCGGTCGCAAA 1137
DB      1081 ACCCTTGTATTAATTGTCATCACTTTTGGGCACTTTAAAGACATGCGGTCGCAAA 1140
QY      1138 CCGGAGGAGGTGGGGATGACCTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTGCACAG 1197
DB      1141 CCGAGGAGGAGGTGGGGATGACCTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTGCACAG 1200
QY      1198 TAAATCAATGAGCGCTGACAGAGGGTTGCCAACCCGAGAGGGAGCTAAATTCAGAAAC 1257
DB      1201 TAAATCAATGAGCGCTGACAGAGGGTTGCCAACCCGAGAGGGAGCTAAATTCAGAAAC 1260
QY      1258 GCGTGTAGTCCGGATCGAGTCTGCACTTCACTTCCGTAAGTGGGAATCCGTAGTAAT 1317
DB      1261 GCGTGTAGTCCGGATCGAGTCTGCACTTCCGTAAGTGGGAATCCGTAGTAAT 1320
QY      1318 CCGGATGACAGATGTCGCGGTGAAATACGTTCCGGGCTTGTACACACCGCCGTACAC 1377
DB      1321 CCGGATGACAGATGTCGCGGTGAAATACGTTCCGGGCTTGTACACACCGCCGTACAC 1380
QY      1378 CATGGAGTGGGTTTCCAGGAAGCAGTATGTTAACTGTAAGAGAGGCGCTTGGCCACGG 1437
DB      1381 CATGGAGTGGGTTTCCAGGAAGCAGTATGTTAACTGTAAGAGAGGCGCTTGGCCACGG 1440
QY      1438 TGAATTCATGACTGGGGTGG 1457
DB      1441 TGAATTCATGACTGGGGTGG 1460

RESULT 59
AY631271
LOCUS      AY631271      1458 bp      DNA      linear      BCT 28-DEC-2004
DEFINITION Nitrosospiira sp. KAN6 16S ribosomal RNA gene, partial sequence.
ACCESSION  AY631271
VERSION     AY631271.1 GI:51093366
KEYWORDS
SOURCE      Nitrosospiira sp. KAN6
ORGANISM    Nitrosospiira sp. KAN6
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosospiira.
REFERENCE   1 (bases 1 to 1458)
AUTHORS     Ida,T., Satoh,M., Yabe,R., Takahashi,R. and Tokuyama,T.
TITLE       Identification of Genus Nitrosospiiro, Ammonia-Oxidizing Bacteria,
by Comparison of N-Terminal Amino Acid Sequences of
Phosphoglycerate Kinase
J. Biosci. Bioeng. 98 (5), 380-383 (2004)
2 (bases 1 to 1458)
AUTHORS     Takahashi,R., Shintaku,Y. and Tokuyama,T.
TITLE       Nitrosolobus sp. KAN6 16S ribosomal RNA gene
Unpublished
3 (bases 1 to 1458)
AUTHORS     Takahashi,R., Shintaku,Y. and Tokuyama,T.
TITLE       Direct Submission
JOURNAL     Submitted (20-MAY-2004) Agricultural and Biological Chemistry,
College of Bioresource Sciences, Nihon University, 1866 Kameino,
Fujisawa, Kanagawa 252-8510, Japan
FEATURES
source
1..1458
/organism="Nitrosospiira sp. KAN6"
/mol_type="genomic DNA"
/db_xref="taxon:289541"
/cldone="KAN6"
/note="similar to Nitrosospiira (Nitrosolobus) multiformis"
<1..>1458
```

ORIGIN  
/product="16S ribosomal RNA"

88.0%; Score 1282.8; DB 1; Length 1458;

Query Match Best Local Similarity 93.4%; Pred. No. 4.3e-13; Mismatches 92; Indels 5; Gaps 2; Matches 1363; Conservative 0;

```

QY 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAATCGCAACGGGAGCAGCGATGCTTGCAAT 60
DB 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAATCGCAACGGGAGCAGCGATGCTTGCAAT 58
QY 61 CTGGTGGCGATGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 59 CTGGTGGCGATGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 118
QY 121 ACCGATGAAAGATGCTCTAATACCGCATATCTCTAAGAGAAACAGGGGATGCGAA 180
DB 119 ACCGATGAAAGATGCTCTAATACCGCATATCTCTAAGAGAAACAGGGGATGCGAA 178
QY 181 GACCTGGCGCTTTGAGCGGCGGATGCTGATGATGATGATGATGATGATGATGATGATGAT 240
DB 179 GACCTGGCGCTTTGAGCGGCGGATGCTGATGATGATGATGATGATGATGATGATGATGAT 238
QY 241 CAAGCGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 239 CAAGCGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298
QY 301 GCCGACATCTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 299 GCCGACATCTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 358
QY 361 CAGCAATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 359 CAGCAATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418
QY 421 AAAGGTTACGGTAAATATATCTGTAACCATGACGGTATGACAGAAAGACCGGCTTAC 480
DB 419 AAAGGTTACGGTAAATATATCTGTAACCATGACGGTATGACAGAAAGACCGGCTTAC 478
QY 481 TACGTGCGAGCAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 479 TACGTGCGAGCAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538
QY 541 AAAGGTTACGGTAAATATATCTGTAACCATGACGGTATGACAGAAAGACCGGCTTAC 600
DB 539 AAAGGTTACGGTAAATATATCTGTAACCATGACGGTATGACAGAAAGACCGGCTTAC 598
QY 601 TGCCTTTGAAACTCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 599 TGCCTTTGAAACTCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 658
QY 661 TGAATGCGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 659 TGAATGCGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
QY 721 GACGCTGACGCAAGAAAGCGGAGGAGCAACAGGATGATGATGATGATGATGATGATGATGAT 780
DB 719 GACGCTGACGCAAGAAAGCGGAGGAGCAACAGGATGATGATGATGATGATGATGATGATGAT 778
QY 781 CTAAACGATGATCACTAGTGTGCTTATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 779 CTAAACGATGATCACTAGTGTGCTTATGATGATGATGATGATGATGATGATGATGATGAT 838
QY 841 AGTTGACCGCTGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 839 AGTTGACCGCTGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 898
QY 901 GCACAAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 899 GCACAAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
QY 961 GACATGATGCGAATTTTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017

```

```

DB 959 GACATGATGCCGAAGCCCGCGAGAGTGGGTGAGCCCGAAGGAAACGGCAACAGATG 1018
QY 1018 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
DB 1019 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
QY 1078 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
DB 1079 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138
QY 1138 CCGGAGAAAGTGGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197
DB 1139 CCGGAGAAAGTGGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
QY 1198 TAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257
DB 1199 TAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1258
QY 1258 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
DB 1259 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1318
QY 1318 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377
DB 1319 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1378
QY 1378 CATGGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1437
DB 1379 CATGGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1438
QY 1438 TGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
DB 1439 TGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458

```

RESULT 60  
LOCUS CO796910 1458 bp DNA linear PAT 19-APR-2004  
DEFINITION Sequence 3 from Patent WO2004026772.  
ACCESSION CO796910  
VERSION CO796910.1 GI:46408536  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1. Hovanec T.A.  
AUTHORS Hovanec T.A.  
TITLE Ammonia-oxidizing bacteria and methods of using and detecting  
JOURNAL Patent: WO 2004026772-A 3 01-APR-2004;  
Aquadria Inc. (US)  
FEATURES  
Source location/Qualifiers  
1. 1458  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="A0B Type B R3clones 16S rDNA"

## ORIGIN

Query Match 88.0%; Score 1282.8; DB 6; Length 1458;  
Best Local Similarity 93.4%; Pred. No. 4.3e-13; Mismatches 92; Indels 5; Gaps 2;  
Matches 1363; Conservative 0;

```

QY 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAATCGCAACGGGAGCAGCGATGCTTGCAAT 60
DB 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAATCGCAACGGGAGCAGCGATGCTTGCAAT 58
QY 61 CTGGTGGCGATGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 59 CTGGTGGCGATGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 118
QY 121 ACCGATGAAAGATGCTCTAATACCGCATATCTCTAAGAGAAACAGGGGATGCGAA 180

```



Db 119 AGCACGGAAGGATGCTAATACCGCATATCTCCACGGAAGAAAAGCAGGGATGCA 178  
QY 181 GACCTTGCGCTTTTGGAGCGCCGATGTCGATTAATAGTATGTTGGGTTAAAGGCTAC 240  
Db 179 GACCTTGCGCTTTTGGAGCGCCGATGTCGATTAATAGTATGTTGGGTTAAAGGCTAC 238  
QY 241 CAAGCGACGATCAGTATGTTGTTGAGAGAGACGACGACACTGAGACTGAGACAG 300  
Db 239 CAAGCGACGATCAGTATGTTGTTGAGAGAGACGACGACACTGAGACTGAGACAG 298  
QY 301 GCCCAGACTCTTAACGAGAGGACGACGAGTGGGAAATTTTGAACAATGGCGCGAACCTGATC 360  
Db 299 GCCCAGACTCTTAACGAGAGGACGACGAGTGGGAAATTTTGAACAATGGCGCGAACCTGATC 358  
QY 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGAGA 420  
Db 359 CAGCCATGCGCGCTGAGTGAAGAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGAGA 418  
QY 421 AAAGGTTACGGTTAAATATGTAACCATGACCGTATCCAGAGAGAGACCGGCTAAC 480  
Db 419 AACGGTACGGCTAAATACCGTATCTGACGTTACCGAGAGAGAGACCGGCTAAC 478  
QY 481 TACGTCACAGACCGCGGTAAATACGATGGGTCGAAGCTTAAATCGAATTAATCTGGCGCT 540  
Db 479 TACGTCACAGACCGCGGTAAATACGATGGGTCGAAGCTTAAATCGAATTAATCTGGCGCT 538  
QY 541 AAAGGTTGCGGACGCGGCTTGTAAATGTCAGATGTGAATCCCGGCTTAACTGGGAAT 600  
Db 539 AAAGGTTGCGGACGCGGCTTGTGTAAATGTCAGATGTGAAGCCCGGCTTAACTGGGAAC 598  
QY 601 TGGGTTGAAACTCAAAAGCTAGATGTTGGCAGAGGAGGTGGAAATTCATGTTGAGAG 660  
Db 599 TGGGTTGAAACTCAAAAGCTAGATGTTGGCAGAGGAGGTGGAAATTCATGTTGAGAG 658  
QY 661 TGAATATGCTAGATATGAGAGAACATCGATGGCGAGCGACCTCTCGGTTAACTACT 720  
Db 659 TGAATATGCTAGATATGAGAGAACATCGATGGCGAGCGACCTCTCGGTTAACTACT 718  
QY 721 GACGCTCATGACAGAAAGCTGGGAGCAAAACGATTAATACCTGGTATGTCACGCC 780  
Db 719 GACGCTCATGACAGAAAGCTGGGAGCAAAACGATTAATACCTGGTATGTCACGCC 778  
QY 781 CTAAACGATGTCAACTAGTTGTTGGGCTTAATTAAGGCTTGTAAAGCTTAACGCGGTGA 840  
Db 779 CTAAACGATGTCAACTAGTTGTTGGGCTTAAACGATGTTGTAAACGAGCTTAACGCGGTGA 838  
QY 841 AGTTGACCGGCTGGGAGATACGCTCGCAAGATTAAACTCAAGAAATGAACGGGAGACC 900  
Db 839 AGTTGACCGGCTGGGAGATACGCTCGCAAGATTAAACTCAAGAAATGAACGGGAGACC 898  
QY 901 GCAACAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAACCTTAACCTACCTT 960  
Db 899 GCAACAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAACCTTAACCTACCTT 958  
QY 961 GACATGATGGAATTTTCTAGAGATAGATTAGTGC---TTGGGAAACGCTAACACAGGTG 1017  
Db 959 GACATGATGGAATTTTCTAGAGATAGATTAGTGC---TTGGGAAACGCTAACACAGGTG 1018  
QY 1018 CTGCAATGAGCTGTCGATGAGCTCGTGTGATGATGTTGGTTAAAGTCCCGCAACGAGCGCA 1077  
Db 1019 CTGCAATGAGCTGTCGATGAGCTCGTGTGATGATGTTGGTTAAAGTCCCGCAACGAGCGCA 1078  
QY 1078 ACCCTTGTCAATTAATGCAATCTTTGGTTGGGCACTTTAATGAGACTGCGGTTGACAA 1137  
Db 1079 ACCCTTGTCAATTAATGCAATCTTTGGTTGGGCACTTTAATGAGACTGCGGTTGACAA 1138  
QY 1138 CCGGAGAGAGGTGGGAGATGACGTCAGATGCTCATGGCCCTTAATGGGAGGAGCTTCAACAG 1197  
Db 1139 CCGGAGAGAGGTGGGAGATGACGTCAGATGCTCATGGCCCTTAATGGGAGGAGCTTCAACAG 1198  
QY 1198 TAAATCAATGCGCGCTTACAGAGGTTGCCAAACCGCGAGGGAGGCTTAATCTCAGAAAGC 1257  
Db 1199 TAAATCAATGCGCGCTTACAGAGGTTGCCAAACCGCGAGGGAGGCTTAATCTCAGAAAGC 1258

QY 1258 GGGTGTAGTCCGGATCCGAGTCTGCAACTGACTCCGTAAGTGGGAATCGCTAGTAAT 1317  
Db 1259 GGGTGTAGTCCGGATCCGAGTCTGCAACTGACTCCGTAAGTGGGAATCGCTAGTAAT 1318  
QY 1318 CCGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGTCACAC 1377  
Db 1319 CCGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGTCACAC 1378  
QY 1378 CATGGAGTGGGTTTCAACAGAGAGATGTCCTAACGTAAGAGAGGCGCTTGCACAG 1437  
Db 1379 CATGGAGTGGGTTTCAACAGAGAGATGTCCTAACGTAAGAGAGGCGCTTGCACAG 1438  
QY 1438 TGAATTCATGACTGGGGTGG 1457  
Db 1439 TGAATTCATGACTGGGGTGG 1458  
RESULT 61  
CS089156  
LOCUS CS089156 1458 bp DNA linear PAT 25-MAY-2005  
DEFINITION Sequence 3 from Patent EP1502948.  
ACCESSION CS089156  
VERSION CS089156.1 GI:66714440  
KEYWORDS  
SOURCE .  
ORGANISM unidentified  
1 unidentified  
unclassified.  
REFERENCE  
AUTHORS Hovanec,T.A. and Burrell,P.C.  
TITLE Ammonia-oxidizing bacteria  
JOURNAL Patent: EP 1502948-A 3 02-FEB-2005;  
Aguaria Inc. (US)  
FEATURES  
source  
1..1458  
Location/Qualifiers  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="Description of Unknown Organism: AOB Type B  
ammonia-oxidizing bacterium represented by R3 clones"  
ORIGIN  
Query Match 88.0%; Score 1282.8; DB 6; Length 1458;  
Best Local Similarity 93.4%; Pred. No. 4.3e-13;  
Matches 1363; Conservative 0; Mismatches 92; Indels 5; Gaps 2;  
QY 1 ATTGAACGCTGCGCGCATCTTTAACAATGACATGACGAGTGAACCGCAGCAGCGATGCTGCAT 60  
Db 1 ATTGAACGCTGCGCGCATCTTTAACAATGACATGACGAGTGAACCGCAGCAGCGATGCTGCAT 58  
QY 61 CTGGTGGCAGTGGGCGAACCGGCTGATTAATGATGGGAACGTAATCAGAAAGAGGGGGATTA 120  
Db 59 CTGGTGGCAGTGGGCGAACCGGCTGATTAATGATGGGAACGTAATCAGAAAGAGGGGGATTA 118  
QY 121 ACGCATCGAAAGATGCTAATACGCAATATCTCTTAAGAGAGAAAGCAGGGGATCGAAA 180  
Db 119 ACGCATCGAAAGATGCTAATACGCAATATCTCTCAACGAGAGAAAGCAGGGGATCGCAA 178  
QY 181 GACCTTGCGCTTTTGGAGCGGCGATGTCGATTAATGATGTTGGTGGGTTAAAGGCTTAC 240  
Db 179 GACCTTGCGCTTTTGGAGCGGCGATGTCGATTAATGATGTTGGTGGGTTAAAGGCTTAC 238  
QY 241 CAAGCGACGATCAGTATGTTGTTGAGAGAGACGACGACACTGAGACTGAGACAG 300  
Db 239 CAAGCGACGATCAGTATGTTGTTGAGAGAGACGACGACACTGAGACTGAGACAG 298  
QY 301 GCCCAGACTCTTAACGAGAGGACGACGAGTGGGAAATTTTGAACAATGGCGCGAACCTGATC 360  
Db 299 GCCCAGACTCTTAACGAGAGGACGACGAGTGGGAAATTTTGAACAATGGCGCGAACCTGATC 358  
QY 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGAGA 420  
Db 359 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGAGA 418



QY 421 AAAGTTACGGTAAATATGTCGATCCCATGACGGTATCGACAGAGAAGACCGGCTAAC 480  
DB 419 AACGGTACGGCTAAATACCGTACCTACGATCCGGAAGAGAAGACCGGCTAAC 478  
QY 481 TACGTCCAGACCGCGGCTAATAGTAGGCTCAAGCCCTTAATCCGAATTACTGGGCGT 540  
DB 479 TACGTCCAGACCGCGGCTAATAGTAGGCTCAAGCCCTTAATCCGAATTACTGGGCGT 538  
QY 541 AAAGGTGGCGACGGCGCTTGAAGTCAGATGTAATCCCGGGCTTAACCTGGGAAT 600  
DB 539 AAAGCGTGGCGACGGCTTGTGAAGTCAGATGTAATCCCGGGCTTAACCTGGGAAT 598  
QY 601 TGGCTTTGAAACTCAAAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTACAG 660  
DB 599 TGGCTTTGAAACTCAAAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTACAG 658  
QY 661 TGAATGCTGATAGATATGGAAGAACTCATGAGCGACAGCCCTCTGGGTTAACACT 720  
DB 659 TGAATGCTGATAGATATGGAAGAACTCATGAGCGACAGCCCTCTGGGTTAACACT 718  
QY 721 GAAGCTCAGACGAAAGCGTGGGAGCAAAAGGATTAGTACCCCTGTAGTCAAGCC 780  
DB 719 GACGCTCAGACGAAAGCGTGGGAGCAAAAGGATTAGTACCCCTGTAGTCAAGCC 778  
QY 781 CTAACGATGTCACTAGTGTGTGGGCTTATTAGGCTTGTAAAGCTAAAGCTGACGCTGA 840  
DB 779 CTAACGATGTCACTAGTGTGTGGGCTTATTAGGCTTGTAAAGCTAAAGCTGACGCTGA 838  
QY 841 AGTTGACCGCTGGGAGATACGGTCCGAAATTAATACTCAAGAAATTGAACGGGAGCC 900  
DB 839 AGTTGACCGCTGGGAGATACGGTCCGAAATTAATACTCAAGAAATTGAACGGGAGCC 898  
QY 901 GCACAAAGCGTGTATATGTGATTAATTCAGTCAACGCAAAACCTTAACCTACCTT 960  
DB 899 GCACAAAGCGTGTATATGTGATTAATTCAGTCAACGCAAAACCTTAACCTACCTT 958  
QY 961 GACATGTAGGGAATTTTCTAGAGATAGATTAGTGC--TTCCGGGAAGCTTAACAGAGTG 1017  
DB 959 GACATGTAGGGAATTTTCTAGAGATAGATTAGTGC--TTCCGGGAAGCTTAACAGAGTG 1018  
QY 1018 CTGATGAGCTGTCTGACGCTGTGTGTGTGAGATGTTGGGTTAACTCCCGCAACGAGCGCA 1077  
DB 1019 CTGATGAGCTGTCTGACGCTGTGTGTGTGAGATGTTGGGTTAACTCCCGCAACGAGCGCA 1078  
QY 1078 ACCCTGTCAATTAATGTCATCACTTTGGTGGGCACTTTAATGAGCTGCGGAGCAAA 1137  
DB 1079 ACCCTGTCAATTAATGTCATCACTTTGGTGGGCACTTTAATGAGCTGCGGAGCAAA 1138  
QY 1138 CCGGAGGAAGGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGAGGCTTCAACAG 1197  
DB 1139 CCGGAGGAAGGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGAGGCTTCAACAG 1198  
QY 1198 TAATACATAGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAAGAAC 1257  
DB 1199 TAATACATAGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAAGAAC 1258  
QY 1258 GCGTGTAGTCCGATCGGAATCTGCAACTGCACTCCGTGAAGTCCGAATGCTAGTAT 1317  
DB 1259 GCGTGTAGTCCGATCGGAATCTGCAACTGCACTCCGTGAAGTCCGAATGCTAGTAT 1318  
QY 1318 CCGGATCAGCATGTCCGCTGAATAGTTCGCGGCTTGTATACACACCGCGCTACAC 1377  
DB 1319 CCGGATCAGCATGTCCGCTGAATAGTTCGCGGCTTGTATACACACCGCGCTACAC 1378  
QY 1378 CATGGAGTGGGTTTACACAGAGCAGTAGTCTTAACGTTAAGAGGCGCTTGCACGG 1437  
DB 1379 CATGGAGTGGGTTTACACAGAGCAGTAGTCTTAACGTTAAGAGGCGCTTGCACGG 1438  
QY 1438 TGAATTCATGACTGGGCTG 1457  
DB 1439 TGAATTCATGACTGGGCTG 1458

RESULT 62  
AXJ16094 1458 bp DNA linear PAT 14-DEC-2001  
LOCUS AXJ16094  
DEFINITION Sequence 3 from Patent WO0190312.  
ACCESSION AXJ16094  
VERSION AXJ16094.1 GI:17899285  
KEYWORDS  
SOURCE  
ORGANISM  
unclassified  
unclassified sequences.  
REFERENCE  
1 Hovanec, T.A. and Burrell, P.C.  
AUTHORS Ammonia-oxidizing bacteria  
TITLE Patent: WO 0190312-A 3 29-NOV-2001;  
JOURNAL AQUARIA, INC. (US)  
FEATURES  
source  
1..1458  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="AOB Type B ammonia-oxidizing bacterium represented  
by R3 clones"

ORIGIN  
Query Match 88.0%; Score 1282.8; DB 6; Length 1458;  
Best Local Similarity 93.4%; Pred. No. 4,3e-13;  
Matches 1363; Conservative 0; Mismatches 92; Indels 5; Gaps 2;  
QY 1 ATGGAACGCTGGGGGATGCTTTACATGCAAGTGCAGACGGGACGAGCTTGCAT 60  
DB 1 ATGGAACGCTGGGGGATGCTTTACATGCAAGTGCAGACGGGACGAGCTTGCAT 58  
QY 61 CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGAAAGTATCCAGAAAGCGGGGTA 120  
DB 59 CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGAAAGTATCCAGAAAGCGGGGTA 118  
QY 121 AGCGATCGAAAGATGTGCTAATACCGCATATCTCTTAAGAGGAAAGCAGGAGATCGAA 180  
DB 119 AGCGATCGAAAGATGTGCTAATACCGCATATCTCTCAAGGAAAGCAGGAGATCGAA 178  
QY 181 GACCTGGCGCTTTTGAAGGGCGGAGTGTGATTAAGCTAGTGTGGGCTTAAAGGCTTAC 240  
DB 179 GACCTGGCGCTTTTGAAGGGCGGAGTGTGATTAAGCTAGTGTGGGCTTAAAGGCTTAC 238  
QY 241 CAAAGCGACGATGATGATGATCTGAGAGAGCAACGACACCTGGGACTGAGACAG 300  
DB 239 CAAAGCGACGATGATGATGATCTGAGAGAGCAACGACACCTGGGACTGAGACAG 298  
QY 301 GCCCAGACTCTTACGGGAGCGACGATGGGGAATTTTGGCAATGGGCGCAACCTGATC 360  
DB 299 GCCCAGACTCTTACGGGAGCGACGATGGGGAATTTTGGCAATGGGCGCAACCTGATC 358  
QY 361 CAGCAATGCCCGCTGAGTGAAGAGGCTTCCGGTGTAAAGCTTTCAGTCCAGAGA 420  
DB 359 CAGCAATGCCCGCTGAGTGAAGAGGCTTCCGGTGTAAAGCTTTCAGTCCAGAGA 418  
QY 421 AAAGTTACGGTAAATATGTCGATCCCATGACGGTATCGACAGAGAAGACCGGCTAAC 480  
DB 419 AACGGTACGGCTAAATACCGTACCTACGATCCGGAAGAGAAGACCGGCTAAC 478  
QY 481 TACGTCCAGACCGCGGCTAATAGTAGGCTCAAGCCCTTAATCCGAATTACTGGGCGT 540  
DB 479 TACGTCCAGACCGCGGCTAATAGTAGGCTCAAGCCCTTAATCCGAATTACTGGGCGT 538  
QY 541 AAAGGTGGCGACGGCGCTTGAAGTCAGATGTAATCCCGGGCTTAACCTGGGAAT 600  
DB 539 AAAGGTGGCGACGGCGCTTGTGAAGTCAGATGTAATCCCGGGCTTAACCTGGGAAT 598  
QY 601 TGGCTTTGAAACTCAAAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTACAG 660  
DB 599 TGGCTTTGAAACTCAAAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTACAG 658

```
QY 661 TGAATGCGTAGATATGAGAAACATCGATGGCCAGGCGCTCTGGTTAACT 720
DB 659 TGAATGCGTAGATATGAGAAACATCGATGGCCAGGCGCTCTGGTTAACT 718
QY 721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAAGTACCTGTAGTCCAGCC 780
DB 719 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAAGTACCTGTAGTCCAGCC 778
QY 781 CTAAAGATGATCACTAGTTGTTGGGCTTTATAGGCTTGTATACCAAGTAAACGGGTGA 840
DB 779 CTAAAGATGATCACTAGTTGTTGGGCTTTATAGGCTTGTATACCAAGTAAACGGGTGA 838
QY 841 AGTTGACCGCTGGGAGTAGTGGTCCGAAAGATTAACTCAAGAAATTGACGGGAGCC 900
DB 839 AGTTGACCGCTGGGAGTAGTGGTCCGAAAGATTAACTCAAGAAATTGACGGGAGCC 898
QY 901 GCACAAAGCGGTGATTAATGATGATTAATTCGATGCAACGGAAAACTTACCTACCTT 960
DB 899 GCACAAAGCGGTGATTAATGATGATTAATTCGATGCAACGGAAAACTTACCTACCTT 958
QY 961 GACATGTAGGGAATTTTCTAGATATAGATTAGTGC---TTCCGAAACGCTTAACAGGTG 1017
DB 959 GACATGTAGGGAATTTTCTAGATATAGATTAGTGC---TTCCGAAACGCTTAACAGGTG 1018
QY 1018 CTGCAATGAGCTGTGCTGAGCTGTGTGCTGATGATTTAGGTTAGTCCCGCAAGAGCGCA 1077
DB 1019 CTGCAATGAGCTGTGCTGAGCTGTGTGCTGATGATTTAGGTTAGTCCCGCAAGAGCGCA 1078
QY 1078 ACCCTGTGCTAATTAATGCTCATCATTTGTTGGGCACTTTAATGAGACTGCGGTGACAA 1137
DB 1079 ACCCTGTGCTAATTAATGCTCATCATTTGTTGGGCACTTTAATGAGACTGCGGTGACAA 1138
QY 1138 CCGGAGAAAGGTGGGATGACGTCAGTCTCATGTCGCCCTTATGGGTAGGGCTTTCACAG 1197
DB 1139 CCGGAGAAAGGTGGGATGACGTCAGTCTCATGTCGCCCTTATGGGTAGGGCTTTCACAG 1198
QY 1198 TAATACATATGCGCGTACAGAGAGGTTGCCAACCCGAGGGAGGCTAATCTCAGAAAGC 1257
DB 1199 TAATACATATGCGCGTACAGAGAGGTTGCCAACCCGAGGGAGGCTAATCTCAGAAAGC 1258
QY 1258 GCGTCGTAGTCCGGAATCGAGTCTGCAACTCGCTCGTAAGTCCGGAATCGTCTAGTAAT 1317
DB 1259 GCGTCGTAGTCCGGAATCGAGTCTGCAACTCGCTCGTAAGTCCGGAATCGTCTAGTAAT 1318
QY 1318 CCGGATCATGCAATGTGCGCGGTGAATACGTTCCCGGGTCTTTGTACACACCGCCGTCACAC 1377
DB 1319 CCGGATCATGCAATGTGCGCGGTGAATACGTTCCCGGGTCTTTGTACACACCGCCGTCACAC 1378
QY 1378 CATGGAGTGGGTTTACCAAGAGAGGTAGTCTAACCCGTAAGAGAGGCGCTTGCCACGG 1437
DB 1379 CATGGAGTGGGTTTACCAAGAGAGGTAGTCTAACCCGTAAGAGAGGCGCTTGCCACGG 1438
QY 1438 TGAGATTCAATGACTGGGGTG 1457
DB 1439 TGAGATTCAATGACTGGGGTG 1458

RESULT 63
AY123796 1497 bp DNA linear BCT 23-Sep-2003
LOCUS Nitrospira sp. L115 16S ribosomal RNA gene, partial sequence.
DEFINITION AY123796
ACCESSION AY123796.1 GI:24474425
VERSION
KEYWORDS
SOURCE
ORGANISM Nitrospira sp. L115
Nitrospira sp. L115
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrospira.
REFERENCE 1 (bases 1 to 1497)
Purkhof, U., Wagner, M., Timmermann, G., Pommerening-Roesser, A. and
Koops, H. P.
16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial
ammonia-oxidizing isolates: extension of the dataset and proposal
```

```
of a new lineage within the nitrosomonads
Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)
JOURNAL 1310037
PUBMED 2 (bases 1 to 1497)
REFERENCE Purkhof, U., Timmermann, G., Koops, H.-P., Pommerening-Roesser, A. and
AUTHORS Wagner, M.
TITLE Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische
JOURNAL Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES
source
1. 1497
/organism="Nitrospira sp. L115"
/mol_type="genomic DNA"
/strain="L115"
/db_xref="taxon:174928"
<1..>1497
/product="16S ribosomal RNA"

ORIGIN
Query Match 88.0%; Score 1282.8; DB 1; Length 1497;
Best Local Similarity 93.4%; Pred. No. 4,2e-13;
Matches 1363; Conservative 0; Mismatches 92; Indels 5; Gaps 2;

QY 1 ATTGAACGCTGGCGGATCTTTACACATGCAAGTCGAACGGCAGCAGCGATGCTTGCAT 60
DB 3 ATTGAACGCTGGCGGATCTTTACACATGCAAGTCGAACGGCAGCAGCGGGAAC-- 60
QY 61 CTGGTGGCAGTGGCGGACGGGTGAGTAATGATGGAACGTAATCCAGAAAGGGGGGTA 120
DB 61 CTGGTGGCAGTGGCGGACGGGTGAGTAATGATGGAACGTAATCCAGAAAGGGGGGTA 120
QY 121 ACGCATCGAAAGATGAGTATACCGCATATCTCTAAGAGAGAAAGCAGGGATCGAA 180
DB 121 ACGCATCGAAAGATGAGTATACCGCATATCTCTAAGAGAGAAAGCAGGGATCGAA 180
QY 181 GACCTTGGCTTTTGAAGCGCCGATGTCATTAAGTAACTTGTGGGTAAAGGCTTAC 240
DB 181 GACCTTGGCTTTTGAAGCGCCGATGTCATTAAGTAACTTGTGGGTAAAGGCTTAC 240
QY 241 CAAAGCGACATCATGATGTTGGTCTGAGAGACACACAGCCACTGCGGACTGAGACAG 300
DB 241 CAAAGCGACATCATGATGTTGGTCTGAGAGACACACAGCCACTGCGGACTGAGACAG 300
QY 301 GCGCAAGCTCTTACGGGAGGACAGGTGGGAAATTTTGGACAAATGGGGAACCTGATC 360
DB 301 GCGCAAGCTCTTACGGGAGGACAGGTGGGAAATTTTGGACAAATGGGGAACCTGATC 360
QY 361 CAGCAATGCGCGTGAATGTAAGAGAGGCTTCCGGTGTAAAGCTCTTTCAGTGAAGAA 420
DB 361 CAGCAATGCGCGTGAATGTAAGAGAGGCTTCCGGTGTAAAGCTCTTTCAGTGAAGAA 420
QY 421 AAAGGTTACGGTAATTAATCGTACCATGACGATGATGCAAGAAAGACACCGGCTAAC 480
DB 421 AAAGGTTACGGTAATTAATCGTACCATGACGATGATGCAAGAAAGACACCGGCTAAC 480
QY 481 TAGGTGACAGACCGCGGTAACTAGTAGGGGTAAGGCTTAATGGAATTAATCTGGGCGT 540
DB 481 TAGGTGACAGACCGCGGTAACTAGTAGGGGTAAGGCTTAATGGAATTAATCTGGGCGT 540
QY 541 AAAGGTTGCGAGCGGCGCTTGAATGATGATGTAATCCCGGCTTAACTGGGAAT 600
DB 541 AAAGGTTGCGAGCGGCGCTTGAATGATGATGTAATCCCGGCTTAACTGGGAAT 600
QY 601 TCGCTTTGAACTCAAAAGCTAGAGTGGGACAGAGGAGGTGGAATTCATGTGACAG 660
DB 601 TCGCTTTGAACTCAAAAGCTAGAGTGGGACAGAGGAGGTGGAATTCATGTGACAG 660
QY 661 TGAATGCTAGATATGAGAAAGATCGATGAGGGAAGGCGCTCCCTGGGTTAACT 720
DB 661 TGAATGCTAGATATGAGAAAGATCGATGAGGGAAGGCGCTCCCTGGGTTAACT 720
QY 721 GACGCTCATGCAAGAGCGTGGGAGCAAAACAGATTAAGTACCTGTAGTCCAGCC 780
```

Db	721	GACGCTCAGGACGAAAGCTGGGGAGCAAAACAGATTAGTACCTTGTAATGCCACGCC	780
Qy	781	CTAAACGATGCTCACTAGTTGTTGGGCTTATTAAGCTTGTAACGAAAGCTAACCGCGTGA	840
Db	781	CTAAACGATGCTCACTAGTTGTTGGGCTTATTAAGCTTGTAACGAAAGCTAACCGCGTGA	840
Qy	841	AGTTGACCGCTTGGGGAGTACCGGTGCGAAGATTAACTCAAGGAATTGACGGGACCC	900
Db	841	AGTTGACCGCTTGGGGAGTACCGGTGCGAAGATTAACTCAAGGAATTGACGGGACCC	900
Qy	901	GCAACAGCGGTGATTTATGTGAAATTAATTCGATGCAACGGGAAACCTTACCTACCTT	960
Db	901	GCAACAGCGGTGATTTATGTGAAATTAATTCGATGCAACGGGAAACCTTACCTACCTT	960
Qy	961	GACATGAGCGAAATTTCTAGAGATAGATTAGTGC--TTCCGGGAGCGTTAACACAGAGTG	1017
Db	961	GACATGAGCGAAATTTCTAGAGATAGATTAGTGC--TTCCGGGAGCGTTAACACAGAGTG	1020
Qy	1018	CTGCATGAGCTGTGCTGACGCTGTGTGCTGAGATGTTGGGTTAACTCCCGCAACGAGCGCA	1077
Db	1021	CTGCATGAGCTGTGCTGACGCTGTGTGCTGAGATGTTGGGTTAACTCCCGCAACGAGCGCA	1080
Qy	1078	ACCTTGTCAATTAATTCGCATCATTTTGGTGGGACCTTTAATGAGACTGCGCGGTGACAA	1137
Db	1081	ACCTTGTCAATTAATTCGCATCATTTTGGTGGGACCTTTAATGAGACTGCGCGGTGACAA	1140
Qy	1138	CCGGAGGAGGTGGGGATGACGTCAGTCTCATGAGCCCTTATGGGTAGGCTTACAGC	1197
Db	1141	CCGGAGGAGGTGGGGATGACGTCAGTCTCATGAGCCCTTATGGGTAGGCTTACAGC	1280
Qy	1198	TAATACATAGCGCGGTACAGAGGGTTGCCAACCCCGAGGGGAGCTAATCTCAAGAAC	1257
Db	1201	TAATACATAGCGCGGTACAGAGGGTTGCCAACCCCGAGGGGAGCTAATCTCAAGAAC	1260
Qy	1258	GCGTCGTATGTCGGATTCGGAATTCGCACTCCGTGAAGTCGGAATGCTAGTAAT	1317
Db	1261	GCGTCGTATGTCGGATTCGGAATTCGCACTCCGTGAAGTCGGAATGCTAGTAAT	1320
Qy	1318	CGCGGATCAGCATGTCGGGTGAATACGTTCCCGGGCTTGTATACACACCCCGCTCAC	1377
Db	1321	CGCGGATCAGCATGTCGGGTGAATACGTTCCCGGGCTTGTATACACACCCCGCTCAC	1380
Qy	1378	CATGGAGTGGGTTTACACAGAACAGAGTACTTAAACCTTAAGAGAGGGCGCTTCCACG	1437
Db	1381	CATGGAGTGGGTTTACACAGAACAGAGTACTTAAACCTTAAGAGAGGGCGCTTCCACG	1440
Qy	1438	TGAGATTCATGACTGGGGTG 1457	
Db	1441	TGAGATTCATGACTGGGGTG 1460	

RESULT 64  
 AY123812 1497 bp DNA linear BCT 23-SEP-2003  
 LOCUS Nitrosospiira sp. N15 16S ribosomal RNA gene, partial sequence.  
 DEFINITION AY123812  
 ACCESSION AY123812.1 GI:24474441  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

Nitrosospiira sp. N15  
 Nitrosospiira sp. N15  
 Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
 Nitrosomonadaceae; Nitrosospiira.  
 1 (bases 1 to 1497)  
 Puthold, U., Wagner, M., Timmermann, G., Pommerening-Roser, A. and  
 Koops, H. P.

TITLE  
 JOURNAL  
 PUBMED  
 REFERENCE  
 AUTHORS

16S rRNA and amoa-based phylogeny of 12 novel betaproteobacterial  
 ammonia-oxidizing isolates: extension of the dataset and proposal  
 of a new lineage within the nitrosomonads  
 Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)  
 2 (bases 1 to 1497)  
 Puthold, U., Timmermann, G., Koops, H.-P., Pommerening-Roser, A. and  
 Wagner, M.

Qy	1	ATTGAACGCTGGGCGCATGCTTTACACATGCAAGTCGAACGGGACGATGCTTGCAAT	60
Db	3	ATTGAACGCTGGGCGCATGCTTTACACATGCAAGTCGAACGGGACGATGCTTGCAAT	60
Qy	61	CTGTGGCGAGTGGCGGACGGGTGAGTATGATCGGAACGTATCCAGAAAGGGGGTGA	120
Db	61	CTGTGGCGAGTGGCGGACGGGTGAGTATGATCGGAACGTATCCAGAAAGGGGGTGA	120
Qy	121	ACGCATCGAAAGATGTCCTAATACCCGATATCTCTTAAGAGAGAAAGCAGGGGATCGAA	180
Db	121	ACGCATCGAAAGATGTCCTAATACCCGATATCTCTTAAGAGAGAAAGCAGGGGATCGAA	180
Qy	181	GACCTTGCGCTTTTGGAGCGGCGGATGTCATGATTAGCTAGTGGGTAAAGCCCTAC	240
Db	181	GACCTTGCGCTTTTGGAGCGGCGGATGTCATGATTAGCTAGTGGGTAAAGCCCTAC	240
Qy	241	CAAGCGCAGCATAGTATGTTGTCAGAGAGCAGCAGCAGCAGCTGGACTGAGACAG	300
Db	241	CAAGCGCAGCATAGTATGTTGTCAGAGAGCAGCAGCAGCAGCAGCTGGACTGAGACAG	300
Qy	301	GCCGACCTCTTACGAGGAGCAGCAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC	360
Db	301	GCCGACCTCTTACGAGGAGCAGCAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC	360
Qy	361	CAGCAATGCGCGGATGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTCGAGAGA	420
Db	361	CAGCAATGCGCGGATGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTCGAGAGA	420
Qy	421	AAAGTTACGTTAATTAATCGTGAACCATGACGGTATGACAGAGAAAGACCGGCTAAC	480
Db	421	AAAGTTACGTTAATTAATCGTGAACCATGACGGTATGACAGAGAAAGACCGGCTAAC	480
Qy	481	TACGTGCGACAGCGCGGTATATCGTAGGTGCAACGCTTAATCGGAATTAATCTGGCGCT	540
Db	481	TACGTGCGACAGCGCGGTATATCGTAGGTGCAACGCTTAATCGGAATTAATCTGGCGCT	540
Qy	541	AAAGGTCGCGAGCGGCTTTGTAATGCAATGGAAGCCCGGCTTAATCTGGGAAAC	600
Db	541	AAAGGTCGCGAGCGGCTTTGTAATGCAATGGAAGCCCGGCTTAATCTGGGAAAC	600
Qy	601	TGCGTTTGAACCTCAAGCTAGAGTGGGAGAGGAGGTGAAATTCATGTTAGACG	660
Db	601	TGCGTTTGAACCTCAAGCTAGAGTGGGAGAGGAGGTGAAATTCATGTTAGACG	660
Qy	661	TGAATGCGTGAAGATGAGAGAACCCGATGCGAAGGACGCCCTTGAGTTAACACT	720
Db	661	TGAATGCGTGAAGATGAGAGAACCCGATGCGAAGGACGCCCTTGAGTTAACACT	720
Qy	721	GACGCTCAGGACGAAAGCGTGGGAGCAAAACAGATTAGTACCTTGTAATGCCACGCC	780
Db	721	GACGCTCAGGACGAAAGCGTGGGAGCAAAACAGATTAGTACCTTGTAATGCCACGCC	780
Qy	781	CTAAACGATGCTCACTAGTTGTTGGGCTTATTAAGCTTGTAACGAAAGCTAACCGCGTGA	840
Db	781	CTAAACGATGCTCACTAGTTGTTGGGCTTATTAAGCTTGTAACGAAAGCTAACCGCGTGA	840

TITLE Direct Submission  
 JOURNAL Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische  
 Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany  
 FEATURES  
 source  
 1. 1497  
 /organism="Nitrosospiira sp. N15"  
 /mol\_type="genomic DNA"  
 /strain="N15"  
 /db\_xref="taxon:200120"  
 <1. >1497  
 /product="16S ribosomal RNA"

ORIGIN  
 Query Match 88.0%; Score 1282.8; DB 1; Length 1497;  
 Best Local Similarity 93.4%; Pred. No. 4,2e-13;  
 Matches 1363; Conservative 0; Mismatches 92; Indels 5; Gaps 2;

QY 841 AGTTGACCGGCTGGGGAGTAGCGGTCGCAAGATTAAAACTCAAGGAATTGACGGGGACCC 900  
| | | | |  
Db 841 AGTTGACCGGCTGGGGAGTAGCGGTCGCAAGATTAAAACTCAAGGAATTGACGGGGACCC 900  
QY 901 GCACAAAGCGGTGGATTATGTGATTATTTGATGCAACGGGAAAACTTTACCTTACCTT 960  
| | | | |  
Db 901 GCACAAAGCGGTGGATTATGTGATTATTTGATGCAACGGGAAAACTTTACCTTACCTT 960  
QY 961 GACATGTAGCGAATTTTCTAGAGTAGATTAGTGC---TTCCGGAAAGCTTAACAGAGTG 1017  
| | | | |  
Db 961 GACATGTAGCGAAGCCCGGAGAGGTGGGTGCCGAAAGGAGCGGTAAACAGAGTG 1020  
QY 1018 CTGCATAGCTGTGCTGACGCTGCTGCTGATGATTTGGTTAAGTCCCGCAAGAGCGCA 1077  
| | | | |  
Db 1021 CTGCATAGCTGTGCTGACGCTGCTGCTGATGATTTGGTTAAGTCCCGCAAGAGCGCA 1080  
QY 1078 ACCCTTGTCAATTAATTTGATGCAATTTGGTTGGGCACTTTAATGAGACTGCCGGTGA 1137  
| | | | |  
Db 1081 ACCCTTGTCAATTAATTTGATGCAATTTGGTTGGGCACTTTAATGAGACTGCCGGTGA 1140  
QY 1138 CCGGAGGAAGGTGGGGATGACGTCAGTCTCTCATGCGCTTATGGGTAGGGCTTCAACG 1197  
| | | | |  
Db 1141 CCGGAGGAAGGTGGGGATGACGTCAGTCTCTCATGCGCTTATGGGTAGGGCTTCAACG 1200  
QY 1198 TAATACAAATGCGCGTACAGAGGATTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAG 1257  
| | | | |  
Db 1201 TAATACAAATGCGCGTACAGAGGATTGCCAACCCGCGAGGGGAGCTAATCTCAGAAAG 1260  
QY 1258 GCGTCGTAGTCCGGATCGAGTCTGCACTGCACTCCGTAAGTCGGAATCGCTAGTAAT 1317  
| | | | |  
Db 1261 GCGTCGTAGTCCGGATCGAGTCTGCACTGCACTCCGTAAGTCGGAATCGCTAGTAAT 1320  
QY 1318 CCGGATTCAGCAGTGGCGGTGAATACGTTCCCGGGCTTTGACACACCGCCGCTACAC 1377  
| | | | |  
Db 1321 CCGGATTCAGCAGTGGCGGTGAATACGTTCCCGGGCTTTGACACACCGCCGCTACAC 1380  
QY 1378 CATGGAGTGGGTTTACCAAGAGCAGTAGTCTAAACCGTAAGAGAGGCGCTTGCCACG 1437  
| | | | |  
Db 1381 CATGGAGTGGGTTTACCAAGAGCAGTAGTCTAAACCGTAAGAGAGGCGCTTGCCACG 1440  
QY 1438 TGAGATTCACTGCTGGGGTG 1457  
| | | | |  
Db 1441 TGAGATTCACTGCTGGGGTG 1460  
RESULT 65  
AF386754 1486 bp DNA linear ENV 10-MAY-2004  
LOCUS AF386754  
DEFINITION Nitrosopila sp. BFl6c46 16S ribosomal RNA gene, partial sequence.  
ACCESSION AF386754  
VERSION AF386754.1 GI:117864831  
KEYWORDS ENV.  
SOURCE Nitrosopila sp. BFl6c46  
ORGANISM Nitrosopila sp. BFl6c46  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosopila; environmental samples.  
REFERENCE 1 (bases 1 to 1486)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Identification of bacteria responsible for ammonia oxidation in  
freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 11722936  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submersion  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
Group, 6100 Condon Dr., Moorpark, CA 93021, USA  
FEATURES  
source location/Qualifiers  
1..1486  
/organism="Nitrosopila sp. BFl6c46"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichments"

ORIGIN  
rRNA  
/db xref="taxon:180513"  
/clone="BFl6c46"  
/environmental\_sample  
<1..>1486  
/product="16S ribosomal RNA"  
Query Match 87.9%; Score 1281.2; DB 3; Length 1486;  
Best Local Similarity 93.3%; Pred. No. 4.4e-13;  
Matches 1362; Conservative 0; Mismatches 93; Indels 5; Gaps 2;  
QY 1 ATTGAACGCTGGCGGATGCTTTAACAATGACAGTGAACGGCAGACGGATGCTTCAT 60  
| | | | |  
Db 16 ATTGAACGCTGGCGGATGCTTTAACAATGACAGTGAACGGCAGACGGGGAAC-- 73  
QY 61 CTGGTGGCAGTGGCGGACGGGTGATTAATGATGGAACGTATCCAGAAAGGGGGGTA 120  
| | | | |  
Db 74 CTGGTGGCAGTGGCGGACGGGTGATTAATGATGGAACGTATCTTCAGATGGGGGATA 133  
QY 121 AGGCATCGAAAGATGCTAATACGCATATCTCTAAGGAGGAAAGCAGGGGATGAAA 180  
| | | | |  
Db 134 AGGCATCGAAAGATGCTAATACGCATATCTCTAAGGAGGAAAGCAGGGGATGCAA 193  
QY 181 GACCTTGCCTTTTGGAGCGCGCATGCTCTAGTTAGCTAGTTGTGGGTAAAGCCTAC 240  
| | | | |  
Db 194 GACCTTGCCTTTTGGAGCGCGCATGCTCTAGTTAGCTAGTTGTGGGTAAAGCCTTAC 253  
QY 241 CAAGCGCAGATCATGATGTTGGTCTGAGAGACGACACGCCACCTGCGACTGACACG 300  
| | | | |  
Db 254 CAAGCGCAGATCATGATGTTGGTCTGAGAGACGACACGCCACCTGCGACTGACACG 313  
QY 301 GCCCAAGTCTTACGGGAGGAGCAGATGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360  
| | | | |  
Db 314 GCCCAAGTCTTACGGGAGGAGCAGATGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 373  
QY 361 CAGCAATGCGCGTGAAGTGAAGAGCCCTCGGGTTGTAAGCTTTTCATGCGAAGA 420  
| | | | |  
Db 374 CAGCAATGCGCGTGAAGTGAAGAGCCCTCGGGTTGTAAGCTTTTCATGCGAAGA 433  
QY 421 AAAGTTACGGTAATTAATCTGATGCCATGACGCTATGACAGAAAGAACCCGCTAAC 480  
| | | | |  
Db 434 AAAGTTACGGTAATTAATCTGATGCCATGACGCTATGACAGAAAGAACCCGCTAAC 493  
QY 481 TAGGTGCCAGACCGCGGTAAATAGTAGGGTGAAGCGTTAATCGGAATTACTGGGCGT 540  
| | | | |  
Db 494 TAGGTGCCAGACCGCGGTAAATAGTAGGGTGAAGCGGTAAATCGGAATTACTGGGCGT 553  
QY 541 AAAGGTGGCAGGCGGCTTGTAAATGATGATGTAATCCCGGGCTTTAATCTGGGAAT 600  
| | | | |  
Db 554 AAAGGTGGCAGGCGGCTTGTAAATGATGATGTAATCCCGGGCTTTAATCTGGGAAT 613  
QY 601 TCGCTTTGAAATCAAAAGCTAGTGTGGCAGAGAGGTGGAAATTCATGTGTAGCAG 660  
| | | | |  
Db 614 TCGCTTTGAAATCAAAAGCTAGTGTGGCAGAGAGGTGGAAATTCATGTGTAGCAG 673  
QY 661 TGAATATGCTAGATATGAGAAATCATGATGGCGAAGCGCTCTCGGTGTTAACT 720  
| | | | |  
Db 674 TGAATATGCTAGATATGAGAAATCATGATGGCGAAGCGCTCTCGGTGTTAACT 733  
QY 721 GACGCTCATGACGAAGCGGTGGGAGCAACAGAGATTAGATACCTGCTAGTCAACGCG 780  
| | | | |  
Db 734 GACGCTCATGACGAAGCGGTGGGAGCAACAGAGATTAGATACCTGCTAGTCAACGCG 793  
QY 781 CTAAACGATGTCAATGATGTTGGGCGCTTATTAAGCTTTGTTGTAAGAACTAACCGCTGA 840  
| | | | |  
Db 794 CTAAACGATGTCAATGATGTTGGGCGCTTATTAAGCTTTGTTGTAAGAACTAACCGCTGA 853  
QY 841 AGTTACCGGCTGGGAGTAGTGGTGCAGAGATTAAACTCAAGGAATTTGACGGGGAACCC 900  
| | | | |  
Db 854 AGTTACCGGCTGGGAGTAGTGGTGCAGAGATTAAACTCAAGGAATTTGACGGGGAACCC 913  
QY 901 GCACAAAGCGGTGGATTATGTGATTATTTGATGCAACGGGAAAACTTTACCTTACCTT 960  
| | | | |

Db 914 GCACAGCGGTGATTATGATGATTAATTCATGCAACCGGAAAACTTACCTACCTT 973  
Qy 961 GACATGAGCGAATTTTCTAGAGTATGATTAAGTGC---TTGGGAAACGCTAACAGAGTG 1017  
Db 974 GACATGATCCGAAGCCCGCTGAAGAGTGTGCTCCGAAAGGAGCGGTAAACACAGGTG 1033  
Qy 1018 CTGCATGCTGTGCTGACAGTCTGTGCTGATGATGTTGGGTTAAAGTCCCGACGAGCGCA 1077  
Db 1034 CTGCATGCTGTGCTGACAGTCTGTGCTGATGATGTTGGGTTAAAGTCCCGACGAGCGCA 1093  
Qy 1078 ACCCTGTCTAATTAATGCTCAATTTGGTGGGCACTTTAATGAGACTGGCGGTGACAA 1137  
Db 1094 ACCCTGTCTAATTAATGCTCAATTTGGGCACTTTAATGAGACTGGCGGTGACAA 1153  
Qy 1138 CCGGAGGAGGTGGGAGTGAAGTCAAGTCCATGAGCCCTTAATGAGGAGGCTTACACG 1197  
Db 1154 CCGGAGGAGGTGGGAGTGAAGTCAAGTCCATGAGCCCTTAATGAGGAGGCTTACACG 1213  
Qy 1198 TAATACAAATGCGCGTACAGAGGGTTGCCAACCCGCGAGGGAGCTAATCTCAGAAAGC 1257  
Db 1214 TAATACAAATGCGCGTACAGAGGGTTGCCAACCCGCGAGGGAGCTAATCTCAGAAAGC 1273  
Qy 1258 GCGTCTGATGTCGGATGAGAGTCTGCACTCTGATCCGTGAAGTCGGAAATCGTAAAT 1317  
Db 1274 GCGTCTGATGTCGGATGAGAGTCTGCACTCTGATCCGTGAAGTCGGAAATCGTAAAT 1333  
Qy 1318 CCGGAGTACAGATGCGCGGTGAATAGTTCGCGGCTTGTATCACACCGCCCTGACAC 1377  
Db 1334 CCGGAGTACAGATGCGCGGTGAATAGTTCGCGGCTTGTATCACACCGCCCTGACAC 1393  
Qy 1378 CATGGAGTGGGTTTCAACGAAAGCAGATAGTCTTAACCGTAAGGAGGGCGCTTCCACG 1437  
Db 1394 CATGGAGTGGGTTTCAACGAAAGCAGATAGTCTTAACCGTAAGGAGGGCGCTTCCACG 1453  
Qy 1438 TGAGATTCAATGACTGGGCTG 1457  
Db 1454 TGAGATTCAATGACTGGGCTG 1473

RESULT 66  
AF386755 1469 bp DNA linear ENV 10-MAY-2004  
LOCUS Nitrosospiira sp. R5c20 16S ribosomal RNA gene, partial sequence.  
DEFINITION AF386755  
ACCESSION AF386755.1 GI:17864832  
VERSION  
KEYWORDS ENV.  
SOURCE Nitrosospiira sp. R5c20  
ORGANISM Nitrosospiira sp. R5c20  
REFERENCE 1 (bases 1 to 1469) Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
AUTHORS Identification of bacteria responsible for ammonia oxidation in  
TITLE freshwater aquaria  
JOURNAL Appl. Environ. Microbiol. 67 (12), 5791-5800 (2001)  
PUBMED 11722936  
REFERENCE 2 (bases 1 to 1469)  
AUTHORS Burrell, P.C., Phalen, C.M. and Hovanec, T.A.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2001) Aquatic Research Laboratory, The Aquaria  
FEATURES Group, 6100 Condon Dr, Moorpark, CA 93021, USA  
source Location/Qualifiers  
1. 1469  
/organism="Nitrosospiira sp. R5c20"  
/mol\_type="genomic DNA"  
/isolation\_source="ammonia-oxidizing bacteria isolated  
from freshwater aquaria enrichment"  
/db\_xref="taxon:180514"  
/clone="R5c20"  
/environmental\_sample  
/environmental\_sample  
/product="16S ribosomal RNA"

ORIGIN  
rRNA

Query Match 87.9%; Score 1280; DB 3; Length 1469;  
Best Local Similarity 93.5%; Pred. No. 4, 7e-13;  
Matches 1359; Conservative 0; Mismatches 90; Indels 5; Gaps 2;  
Qy 1 ATTGAACGCTGGCGGATCTTTACATGCAAGTGAACGGGACGAGCTTGCAAT 60  
Db 18 ATTGAACGCTGGCGGATCTTTACATGCAAGTGAACGGGACGAGCTTGCAAT 75  
Qy 61 CTGGTGGCAGATGGCGGAGCGGTTAGTAATGCAATCGGAACGTTTCCAGAAAGAGGGGAT 120  
Db 76 CTGGTGGCAGATGGCGGAGCGGTTAGTAATGCAATCGGAACGTTTCCAGAAAGAGGGGAT 135  
Qy 121 ACCGATCGAAAGATGTCCTAATACCGCATATCTCTAAGAGGAAAGCAGGGATCGAA 180  
Db 136 ACCGATCGAAAGATGTCCTAATACCGCATATCTCTAAGAGGAAAGCAGGGATCGAA 195  
Qy 181 GACCTTGCGCTTGGAGCGGCGATGTCGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 240  
Db 196 GACCTTGCGCTTGGAGCGGCGGATGTCGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 255  
Qy 241 CAAGGCGACGATCAGTATGTTGTTCTGAGAGGACGACACGACACTGGGACTGAGACG 300  
Db 256 CAAGGCGACGATCAGTATGTTGTTCTGAGAGGACGACACGACACTGGGACTGAGACG 315  
Qy 301 GCCGACACTCTTACCGGAGGACGACAGTGGGAAATTTTGAACATGGGGCGCAAGCTGATC 360  
Db 316 GCCGACACTCTTACCGGAGGACGACAGTGGGAAATTTTGAACATGGGGCGCAAGCTGATC 375  
Qy 361 CAGCAATGCGCGGTGAAGTGAAGAGGCTTCGGGTTGAAGCTCTTTCAGTGAGAGA 420  
Db 376 CAGCAATGCGCGGTGAAGTGAAGAGGCTTCGGGTTGAAGCTCTTTCAGTGAGAGA 435  
Qy 421 AAAGTTACCGTAAATATGCTGAACCCATGACGGTATCGACAGAGAGCAACCGGCTTAC 480  
Db 436 AAAGTTACCGTAAATATGCTGAACCCATGACGGTATCGACAGAGAGCAACCGGCTTAC 495  
Qy 481 TAGCTGCCAGAGCCCGGTGAATATCGTAGGTTGCAAGCTTAAATTCAGTGGGCGT 540  
Db 496 TAGCTGCCAGAGCCCGGTGAATATCGTAGGTTGCAAGCTTAAATTCAGTGGGCGT 555  
Qy 541 AAAGGTCGCCAGGCGGCTTGTAGTCAAGTGAATCCCGGCTTAACTCGGGAAT 600  
Db 556 AAAGGTCGCCAGGCGGCTTGTAGTCAAGTGAATCCCGGCTTAACTCGGGAAT 615  
Qy 601 TGCGTTTGAACCTACAAAGCTAGAGTGTGCGAGAGGAGTGGATTCATGTGTACAG 660  
Db 616 TGCGTTTGAACCTACAAAGCTAGAGTGTGCGAGAGGAGTGGATTCATGTGTACAG 675  
Qy 661 TGAATTCGTAAGATATGGAAGACATCGATGGCGGAAGCGACGCTTCGGGTTAAACT 720  
Db 676 TGAATTCGTAAGATATGGAAGACATCGATGGCGGAAGCGACGCTTCGGGTTAAACT 735  
Qy 721 GACGCTCATGACGAAGCGGTGGGAGCAACAGATTTAGTACCTGTGTAGTCCAGCC 780  
Db 736 GACGCTCATGACGAAGCGGTGGGAGCAACAGATTTAGTACCTGTGTAGTCCAGCC 795  
Qy 781 CTAACGATGCACTAGTTGTTGGGCTTATTAAGCTTGTGTAAAGCTAACGCGTGA 840  
Db 796 CTAACGATGCACTAGTTGTTGGGCTTATTAAGCTTGTGTAAAGCTAACGCGTGA 855  
Qy 841 AGTTGACCGCTGGGAGTACGCTGTGCAAGATTAACTCAAGAAATTGACGGGAGCC 900  
Db 856 AGTTGACCGCTGGGAGTACGCTGTGCAAGATTAACTCAAGAAATTGACGGGAGCC 915  
Qy 901 GCACAGGCGGTGATTTATGAGATTAATTTGATGCAACGCGAAACCTTACCTACCTT 960  
Db 916 GCACAGGCGGTGATTTATGAGATTAATTTGATGCAACGCGAAACCTTACCTACCTT 975  
Qy 961 GACATGTAGCGAATTTTCTAGAGTATGATGTC---TTGCGGAACGCTTAACAGGTG 1017  
Db 976 GACATGTAGCGAATTTTCTAGAGTATGATGTC---TTGCGGAACGCTTAACAGGTG 1035



QY 1078 ACCCTGTCAATTAATTCATCAATTTGGTGGCACTTTAATGAGCTGCGGTGACAA 1137  
|||||  
DB 1089 ACCCTGTCAATTAATTCATCAATTTGGTGGCACTTTAATGAGCTGCGGTGACAA 1148  
|||||  
QY 1138 CCGAGAGAGGTGGGGATGACGTCAAGTCTCATGCGCCCTTAATGGGATGGCTTTCACAG 1197  
|||||  
DB 1149 CCGAGAGAGGTGGGGATGACGTCAAGTCTCATGCGCCCTTAATGGGATGGCTTTCACAG 1208  
|||||  
QY 1198 TAATACAAATGGCCCGTACAGAGGGTTGGCAACCGCGAGGGGGAGCTAATCTCAGAAAGC 1257  
1209 TAATACAAATGGCCCGTACAGAGGGTTGGCAACCGCGAGGGGGAGCTAATCTCAGAAAGC 1268  
DB  
QY 1258 GCGTCGATGTCGGGATCGAGTCTGCAACTGCACTCCGCTGAAGTCGGAATCGCTAGTAAT 1317  
1269 GCGTCGATGTCGGGATCGAGTCTGCAACTGCACTCCGCTGAAGTCGGAATCGCTAGTAAT 1328  
DB  
QY 1318 CCGGATCAGCATGTCGCGGTGAATACGTTCCGCGTCTTGTACACACCGCCGTCACAC 1377  
1329 CCGGATCAGCATGTCGCGGTGAATACGTTCCGCGTCTTGTACACACCGCCGTCACAC 1388  
DB  
QY 1378 CATGCAATGGGTTTCCACAGAGCAGTAGTCTTAACCGTAAGGAGGCGCTTGCACAG 1437  
1389 CATGCAATGGGTTTCCACAGAGCAGTAGTCTTAACCGTAAGGAGGCGCTTGCACAG 1448  
DB  
QY 1438 TGAGATTCATG 1448  
1449 TGAGATTCATG 1459  
DB

RESULT 68  
AY856375 1458 bp DNA linear BCT 15-JAN-2005  
LOCUS Nitrosospiira sp. HBN8222A 16S ribosomal RNA gene, partial sequence.  
DEFINITION  
ACCESSION AY856375  
VERSION AY856375.1 GI:57338938  
KEYWORDS  
SOURCE Nitrosospiira sp. HBN8222A  
ORGANISM Nitrosospiira sp. HBN8222A  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
Nitrosomonadaceae; Nitrosospiira.  
REFERENCE  
AUTHORS 1 (bases 1 to 1458)  
Takahashi, R., Yabe, R. and Tokuyama, T.  
TITLE Direct Submission  
JOURNAL Submitted (14-DEC-2004) Agricultural and Biological Chemistry,  
College of Bioresource Sciences, Nihon University, 1866 Kameino,  
Fujisawa, Kanagawa 252-8510, Japan  
Location/Qualifiers  
FEATURES  
source  
1..1458  
/organism="Nitrosospiira sp. HBN8222A"  
/mol\_type="genomic DNA"  
/strain="HBN8222A"  
/db\_xref="taxon:308063"  
/note="Similar to Nitrosospiira (Nitrosolobus) multiformis"  
<1..>1458  
/product="16S ribosomal RNA"

ORIGIN  
Query Match 87.7%; Score 1278; DB 1; Length 1458;  
Best Local Similarity 93.2%; Pred. No. 5.1e-13;  
Matches 1360; Conservative 0; Mismatches 95; Indels 5; Gaps 2;

QY 1 ATTTAAGCGCTGGCGGATGCTTTACATGCAAGTCGAACGGCAGCAGGATCTTGAT 60  
DB 1 ATTTAAGCGCTGGCGGATGCTTTACATGCAAGTCGAACGGCAGCAGGAGGCAAC-- 58  
QY CTGATGGAGAGTGGCGGACGGGTGAGTAATGATCGAAGCGTATCCAGAGAGGGGGGTA 120  
DB CTGATGGAGAGTGGCGGACGGGTGAGTAATGATCGAAGCGTATCCAGAGAGGGGGGTA 118  
QY 121 ACGCATCGAAGATGCTTAATACCGCATATCTTAAAGAGAGAAAGCAGGGAGATCGAA 180  
DB 119 ACGCATCGAAGATGCTTAATACCGCATATCTTAAAGAGAGAAAGCAGGGAGATCGCA 178

QY 181 GACCTTGGCCTTTTGGAGGGCGCGAATGCTAGTTAGCTAGTTGGTGGGTTAAAGGCTTAC 240  
179 GACCTTGGCCTTTTGGAGGGCGCGAATGCTAGTTAGCTAGTTGGTGGGTTAAAGGCTTAC 238  
DB  
QY 241 CAAGCGAGATCAAGATTTGGTCTGAGAGAGCAGCAACGACCACTCTGGAGCTTGAGACAG 300  
239 CAAGCGATGATCAAGATTTGGTCTGAGAGAGCAGCAACGACCACTCTGGAGCTTGAGACAG 298  
DB  
QY 301 GCCCAGACTCTTACGGGAGGACAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCTGATC 360  
299 GCCCAGACTCTTACGGGAGGACAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCTGATC 358  
DB  
QY 361 CAGCAATGCCGCGGTGATGAGAGAGGCTTCCGCTGTAAAGCTCTTCAAGTGAGAGA 420  
359 CAGCAATGCCGCGGTGATGAGAGAGGCTTCCGCTGTAAAGCTCTTCAAGTGAGAGA 418  
DB  
QY 421 AAAGTTACGGTTAATTAATGTAACCCATGACGGTATCGACAGAGAGAGCAGCGGCTAAC 480  
419 AAAGTTACGGTTAATTAATGTAACCCATGACGGTATCGACAGAGAGAGCAGCGGCTAAC 478  
DB  
QY 481 TACGTGCAGACCGCGGTAATACGTAGGCTGCAAGCTTAAATCGGAATTAATCTGGGCGT 540  
479 TACGTGCAGACCGCGGTAATACGTAGGCTGCAAGCTTAAATCGGAATTAATCTGGGCGT 538  
DB  
QY 541 AAAGGTCGCGCAGCGCGCTTGTAAATGATGTAATCCCGGCGCTTAACCTGGGAT 600  
539 AAAGGTCGCGCAGCGCGCTTGTAAATGATGTAATCCCGGCGCTTAACCTGGGAT 598  
DB  
QY 601 TGGGTTTGAACCTAACAAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTTAGACAG 660  
599 TGGGTTTGAACCTAACAAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTTAGACAG 658  
DB  
QY 661 TGAATATCGTAGATATGGAAGAACAATCGATGCGAGCAGCTCTCTGGGTTAAACT 720  
659 TGAATATCGTAGATATGGAAGAACAATCGATGCGAGCAGCTCTCTGGGTTAAACT 718  
DB  
QY 721 GACGCTATGACAGAAAGCTGGGGAGCAACAGAAATTAATACCCGTGATGTCACAGCC 780  
719 GACGCTATGACAGAAAGCTGGGGAGCAACAGAAATTAATACCCGTGATGTCACAGCC 778  
DB  
QY 781 CTAAACGATGCAACTAGTTGTTGGGCTTATTAGCTTGTGTAACGAAGCTTAACGCGTGA 840  
779 CTAAACGATGCAACTAGTTGTTGGGCTTATTAGCTTGTGTAACGAAGCTTAACGCGTGA 838  
DB  
QY 841 AGTTGACCGCTGGGAGTAGCGGTGCGAAGATTAAATCTCAAGAAATTGACGCGGACCC 900  
839 AGTTGACCGCTGGGAGTAGCGGTGCGAAGATTAAATCTCAAGAAATTGACGCGGACCC 898  
DB  
QY 901 GCACAGCGGTGATTAATGATTAATTTGATGCAACGCGAAACCTTAACCTACCCCT 960  
899 GCACAGCGGTGATTAATGATTAATTTGATGCAACGCGAAACCTTAACCTACCCCT 958  
DB  
QY 961 GACATGTAAGGAATTTTCTAGAGTAAGTATAGTGC--TTCCGGAGCGCTAACACAGGTG 1017  
959 GACATGTAAGGAATTTTCTAGAGTAAGTATAGTGC--TTCCGGAGCGCTAACACAGGTG 1018  
DB  
QY 1018 CTGCATGCTGTCTGACGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1077  
1019 CTGCATGCTGTCTGACGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1078  
DB  
QY 1078 ACCCTTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
1079 ACCCTTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138  
DB  
QY 1138 CCGAGAGAGGTGGGGATGACGTCAAGTCTCATGCGCCCTTAATGGGATGGCTTTCACAGC 1197  
1139 CCGAGAGAGGTGGGGATGACGTCAAGTCTCATGCGCCCTTAATGGGATGGCTTTCACAGC 1198  
DB  
QY 1198 TAATACATGCGCGTACAGAGGGTTGCCAACCGCGAGGGGAGCTTAATCTCAGAAAGC 1257  
1199 TAATACATGCGCGTACAGAGGGTTGCCAACCGCGAGGGGAGCTTAATCTCAGAAAGC 1258  
DB



QY	1258	GGCTGCTAGTCCGGANTCCGAAGCTCACA	CTCGA	CTCCGGAGATCGGAAATCGCTAGTAT	131
Db	1259	GGGTGTATCCGGATCCGAATCGGAATCTG	CAATCT	CCGGAAATCGCTAGTAT	1318
QY	1318	CGCGATCAGCAATGTCTGCGCGTGAATAC	CGTTCC	CGGGTCTTGTACACACGCCCTGCAAC	1377
Db	1319	CGCGATCAGCAATGTCTGCGCGTGAATAC	CGTTCC	CGGGTCTTGTACACACGCCCTGCAAC	1378
QY	1378	CATGGAGTGGGTTTACCAAGACAGTAGTCT	TAACCG	TAAAGAGCGCTTGCACCG	143
Db	1379	CATGGAGTGGGTTTACCAAGACAGTAGTCT	TAACCG	TAAAGAGAGGCGCTGCACCG	1438
QY	1438	TGAGATTCAATGACTGGGCTG	1457		
Db	1439	TGAGATTCAATGACTGGGCTG	1458		
RESULT 69					
LOCUS	AY123807	1498 bp	DNA	linear	BCT 23-SEP-2003
DEFINITION	Nitrosospiira multiformis 16S ribosomal RNA gene, partial sequence.				
ACCESSION	AY123807				
VERSION	AY123807.1	GI:24474436			
KEYWORDS					
SOURCE					
ORGANISM	Nitrosospiira multiformis ATCC 25196 (Nitrosolobus multiformis)				
REFERENCE	Nitrosospiira multiformis ATCC 25196 Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosospiira.				
AUTHORS	1 (bases 1 to 1498) Purkhold,U., Wagner,M., Timmermann,G., Pommerening-Roese,A. and Koops,H.P.				
TITLE	16S rRNA and amoA-based phylogeny of 12 novel betaproteobacterial ammonia-oxidizing isolates: extension of the dataset and proposal of a new lineage within the nitrosomonads				
JOURNAL	Int. J. Syst. Evol. Microbiol. 53 (Pt 5), 1485-1494 (2003)				
PUBMED	13130037				
REFERENCE	2 (bases 1 to 1498) Purkhold,U., Timmermann,G., Koops,H.-P., Pommerening-Roeser,A. and Wagner,M.				
AUTHORS	Direct Submission Submitted (19-JUN-2002) Lehrstuhl fuer Mikrobiologie, Technische Universitaet Muenchen, Am Hochanger 4, Freising 85350, Germany				
TITLE	Location/Qualifiers				
JOURNAL	1. 1498				
FEATURES	/organism="Nitrosospiira multiformis ATCC 25196" /mol type="genomic DNA" /strain="N113; C-71; ATCC 25196" /db_xref="ATCC:25196" /db_xref="taxon:323848" /note="type strain synonym: Nitrosolobus multiformis" <1..>1498 /product="16S ribosomal RNA"				
ORIGIN					
Query Match	87.5%	Score 1275;	DB 1;	Length 1498;	
Best Local Similarity	93.4%;	Pred. No. 5.5e-13;			
Matches 1365;	Conservative 0;	Mismatches 90;	Indels 6;	Gaps 3;	
QY	1	ATTGAACGCTGGCGGCATGCTTTACATGCATGC	AAAGCGGACGACAGGATGCTTGAT	60	
Db	3	ATTGAACGCTGGCGGCATGCTTTACATGCATGC	AAAGCGGACGACAGGATGCTTGAT	60	
QY	61	CTGGAGGAGATGGGCGGACGGGGTGAATGATCG	AAACGATCAGAAAGAGGGGGAT	120	
Db	61	CTGGAGGAGATGGGCGGACGGGGTGAATGATCG	AAACGATCAGAAAGAGGGGGAT	120	
QY	121	ACGCATCGAAAGATGTGCTAATACCGCATATACT	CTTAAAGAGGAAAGCGGATCGAA	180	
Db	121	ACGCATCGAAAGATGTGCTAATACCGCATATACT	CTTAAAGAGGAAAGCGGATCGCA	180	
QY	181	GACCTTGCGCTTTTGAGCGCGCGCATGTCTGATT	AGCTAGTTGGGGTAAAGGCTTAC	240	

Db	101	GAACCTTGGGCTTTTGGAGCGGCCGCAATGCTGATTAGCTAGTTGGTAGGTAAGGATTAAAGCTTAC	240
Qy	241	CAAGCGCAGATCAGTAGTTGGTCTTGAGAGACGACACGACCAACTGGGACTGAGACACG	300
Db	241	CAAGCGCTCGATCAGTAGCTGGTCTGAGAGAGCAGCACGACCACTGGGACTGAGACACG	300
Qy	301	GGCCAGACTCTTACGGGAGGACGAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC	360
Db	301	GGCCAGACTCTTACGGGAGGACGAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC	360
Qy	361	CAGCAATGCCGCGTGAAGAGAGGCTTCGGGTGTAAAGCTCTTTCACTCGAAGAG	420
Db	361	CAGCATGCGCGGTGAAGAGAGGCTTCGGGTGTAAAGCTCTTTCAAGCTCTTTCAAGCGGAAACGA	420
Qy	421	AAAGGTTACGGTAAATTAATCGTACCCTAGACGCTATCGACAGAGAGACACCGGCTTAC	480
Db	421	AAAGGTTACGGTAAATTAATCGTACCCTAGACGCTATCGACAGAGAGAGACACCGGCTTAC	480
Qy	481	TACGTGCCAGAGCGCCGGTATACGTAGGGGTGCAAGCTTAATCGAAATTAATTCGGGCGT	540
Db	481	TACGTGCCAGAGCGCCGGTAAATACGTAGGGGTGCAAGCTTAATCGAAATTAATTCGGGCGT	540
Qy	541	AAAGGCTGCGCAGGCGG--CCTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAA	599
Db	541	AAAGGCTGCGCAGGCGGTTTTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAA	600
Qy	600	TTGCGTTTGAAACTTACAAAGCTAGAGTGTGGCAGAGGGAGGTGAATTCATGTGTAGCA	659
Db	601	CTGCGTTTGAAACTTACAAAGCTAGAGTGTGGCAGAGGGGGGTGAATTCATGTGTAGCA	660
Qy	660	GTGAATACCGTAGAATATGGAAGAACTCGATGGCGAAGGAGCCTCCTGGGTTAACAC	719
Db	661	GTGAATACCGTAGAATGTGGAGAACACCGATGGCGAAGGAGCCTCCTGGGTTAACAC	720
Qy	720	TGACGCTCATGACGAAAGCGTGGGAGGACAAACAGATTAAGTACCTTGTAAGTCCACGC	779
Db	721	TGACGCTCATGACGAAAGCGTGGGAGGACAAACAGATTAAGTACCTTGTAAGTCCACGC	780
Qy	780	CCTAAACGATGTCACTAGTGTGGGCTTAAATTAAGCTTGGTAAAGAACTAACCGGTG	839
Db	781	CCTAAACGATGTCACTAGTGTGGGCTTAAACGAACTTGGTAAACGAGCTAACCGGTG	840
Qy	840	AAAGTTGACCCGCTGGGAGGTACGGTTCGCAAGTTAAACTCAAGAGATTTGACGGGAGC	899
Db	841	AAAGTTGACCCGCTGGGAGGTACGGTTCGCAAGTTAAACTCAAGAGATTTGACGGGAGC	900
Qy	900	CGCACACGCGGTGAATATGTGGAATTAATTCGATCGAACGCGAAAACTTACCTTACCT	959
Db	901	CGCACACGCGGTGAATATGTGGAATTAATTCGATCGAACGCGAAAACTTACCTTACCT	960
Qy	960	TGACACTGTAGCGAATTTTCTAGAGATPAGATTAGTSC--TTGCGGAAACGCTAACACAGGT	101
Db	961	TGACACTGTAGCGGAGCCCGCGAGAGGTGGGTGTCCCGAAAGGAAACGGCAACACAGGT	102
Qy	1017	GCTGATGGCTGTGCTCAGCTGTGCTGTAAGTTGGGTTAAGTCCCGCAACGAGCGC	107
Db	1021	GCTGATGGCTGTGCTCAGCTGTGCTGTAAGTTGGGTTAAGTCCCGCAACGAGCGC	108
Qy	1077	AAACCTTGTCAATTAATTCGATCATTTGGTGGGACCTTAATGACCTGCGGTGACAA	113
Db	1081	AAACCTTGTCAATTAATTCGATCATTTGAGTTGGGACCTTAATGAACTGCGGTGACAA	114
Qy	1137	ACCGAGAGAGTGGGAGTGAAGTCAAGTCTCATGATCGCTTATGGTAGGGCTTACAC	119
Db	1141	ACCGAGAGAGTGGGAGTGAAGTCTCATGATCGCTTATGGTAGGGCTTACAC	120
Qy	1197	GTAATATCAATGACGCGTACAGAGGTTGCAACCCCGAGAGGGAGCTAATTCAGAAAG	125
Db	1201	GTAATATCAATGACGCGTACAGAGGTTGCAACCCCGAGAGGGAGCTAATTCAGAAAG	126
Qy	1257	CGCGTCGTAAGTCCGATGCGAGTCTGCAACTCGATCCGTGAAGTCCGAAATGCGTAAGTA	131
Db	1261	CGCGTCGTAAGTCCGATGCGAGTCTGCAACTCGATCCGTGAAGTCCGAAATGCGTAAGTA	132



QY 1317 TCGCGATGAGCATGTCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCGTCACA 1376  
| | | | |  
Db 1321 TCGCGATGAGCATGTCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCGTCACA 1380  
| | | | |  
QY 1377 CCATGGAGTGGGTTTCCACCAAGAGATGATCTAACCGGTAAAGAGGGGCTTTGCCACG 1436  
| | | | |  
Db 1381 CCATGGAGTGGGTTTCCACCAAGAGATGATCTAACCGGTAAAGAGGGGCTTTGCCACG 1440  
| | | | |  
QY 1437 GTGAGATTCACTGACGGGGTG 1457  
| | | | |  
Db 1441 GTGAGATTCACTGACGGGGTG 1461  
| | | | |  
RESULT 70  
AY856376 1457 bp DNA linear BCT 15-JAN-2005  
AY856376  
LOCUS Nitrosospira sp. PM2 16S ribosomal RNA gene, partial sequence.  
DEFINITION  
ACCESSION AY856376  
VERSION  
KEYWORDS  
SOURCE Nitrosospira sp. PM2  
ORGANISM Nitrosospira sp. PM2  
Bacteria; Proteobacteria; Betaproteobacteria; Nitrospomonadales;  
Nitrospomonadaceae; Nitrosospira.  
REFERENCE  
AUTHORS 1. (bases 1 to 1457)  
TITLE Takahashi, R., Yabe, R. and Tokuyama, T.  
JOURNAL Submitted (14-DEC-2004) Agricultural and Biological Chemistry,  
College of Bioresource Sciences, Nihon University, 1866 Kameino,  
Fujisawa, Kanagawa 252-8510, Japan  
location/Qualifiers  
FEATURES  
source 1..1457  
/organism="Nitrosospira sp. PM2"  
/mol\_type="genomic DNA"  
/strain="PM2"  
/db\_xref="taxon:308064"  
/note="similar to Nitrosospira (Nitrosolobus) multiformis"  
<1..>1457  
/product="16S ribosomal RNA"  
ORIGIN  
Query Match 87.3%; Score 1272.4; DB 1; Length 1457;  
Best Local Similarity 93.4%; Pred. No. 6,2e-13;  
Matches 1363; Conservative 0; Mismatches 91; Indels 6; Gaps 3;  
QY 1 ATTTGAAGCTGGCGGATGCTTTTACATCGAAGTCCGAACGGGACGCGATGCTTGAT 60  
| | | | |  
Db 1 ATTGAAGCTGGCGGATGCTTTTACATCGAAGTCCGAACGGGACGCGATGCTTGAT 58  
| | | | |  
QY 61 CTGGTGGGAGTGGCGGACGGGTGATGATCATCGAAGCTATTCAGAAAGAGGGGGTA 120  
| | | | |  
Db 59 CTGGTGGGAGTGGCGGACGGGTGATGATCATCGAAGCTATTCAGAAAGAGGGGGTA 118  
| | | | |  
QY 121 ACGCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGAGGGGATCGAA 180  
| | | | |  
Db 119 ACGCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGAGGGGATCGAA 178  
| | | | |  
QY 181 GACCTTGGCTTTTGGAGCGGCGCATGTCTGATTAAGCTAGTGTGGGTAAAGGCGTAC 240  
| | | | |  
Db 179 GACCTTGGCTTTTGGAGCGGCGCATGTCTGATTAAGCTAGTGTGGGTAAAGGCGTAC 238  
| | | | |  
QY 241 CAAAGCGAGCATGATGATGTTGCTGAGAGACGACCACTGGGATGAGACAG 300  
| | | | |  
Db 239 CAAAGCGAGCATGATGATGTTGCTGAGAGACGACCACTGGGATGAGACAG 298  
| | | | |  
QY 301 GCCCAGACTCTTAACGGAGGACGAGAGTGGGAAATTTGACATGAGGCGCAAGCTGATC 360  
| | | | |  
Db 299 GCCCAGACTCTTAACGGAGGACGAGAGTGGGAAATTTGACATGAGGCGCAAGCTGATC 358  
| | | | |  
QY 361 CAGCATGCGCGGTGATGAGAGAGGCTTGGGTTGTTAAAGCTTTTCAATCGAGAGA 420  
| | | | |  
Db 359 CAGCATGCGCGGTGATGAGAGAGGCTTGGGTTGTTAAAGCTTTTCAATCGAGAGA 418  
| | | | |

QY 421 AAAGTTACGTTAAATATATCGTGAACCATGACGGTATCGACAGAAAGACACGGGCTAAC 480  
| | | | |  
Db 419 AAAGTTACGTTAAATATATCGTGAACCATGACGGTATCGACAGAAAGACACGGGCTAAC 478  
| | | | |  
QY 481 TACGTCCAGACAGCCCGGTTAATACGTAAGGATGCAAGCGTTAATCGAAATTAATGAGCGT 540  
| | | | |  
Db 479 TACGTCCAGACAGCCCGGTTAATACGTAAGGATGCAAGCGTTAATCGAAATTAATGAGCGT 538  
| | | | |  
QY 541 AAAGGTGCGCAGCGCGCTTGTAGTCAAGTGTAAATCCCGGCTTAACCTGGGAAT 600  
| | | | |  
Db 539 AAAGGTGCGCAGCGCGCTTGTAGTCAAGTGTAAATCCCGGCTTAACCTGGGAAT 598  
| | | | |  
QY 601 TCGGTTGAAACTCAACAAGCTGAGTGTGGCAGAGGAGGTGAATTCATGTGTAGAG 660  
| | | | |  
Db 599 TCGGTTGAAACTCAACAAGCTGAGTGTGGCAGAGGAGGTGAATTCATGTGTAGAG 658  
| | | | |  
QY 661 TGAATGCGTAGAGATATGAGAGAACATGATGATGCGAAGGACGCTCTGGGTTAACT 720  
| | | | |  
Db 659 TGAATGCGTAGAGATATGAGAGAACATGATGATGCGAAGGACGCTCTGGGTTAACT 717  
| | | | |  
QY 721 GACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTGATACCTGTTAGTCCAGCC 780  
| | | | |  
Db 718 GACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTGATACCTGTTAGTCCAGCC 777  
| | | | |  
QY 781 CTAAAGCATGTCATTAATGTTGGGCTTAAAGCTTGTATACGAAAGCTAAACGCGTGA 840  
| | | | |  
Db 778 CTAAAGCATGTCATTAATGTTGGGCTTAAAGCTTGTATACGAAAGCTAAACGCGTGA 837  
| | | | |  
QY 841 AGTTGACCGCTGGGAGTACGCTGCGCAAGATTAAACTCAAGAAATTGACGGGAGCC 900  
| | | | |  
Db 838 AGTTGACCGCTGGGAGTACGCTGCGCAAGATTAAACTCAAGAAATTGACGGGAGCC 897  
| | | | |  
QY 901 GCAAGAGCGGTGATTAATGATTAATTCATGCAACCGGAAACCTTAACCTACCTT 960  
| | | | |  
Db 898 GCAAGAGCGGTGATTAATGATTAATTCATGCAACCGGAAACCTTAACCTACCTT 957  
| | | | |  
QY 961 GACATGTAAGCAATTTTCTGAAGATGATTAATGCTGCTGCGGAGCGTTAAACAGAGTG 1017  
| | | | |  
Db 958 GACATGTAAGCAATTTTCTGAAGATGATTAATGCTGCTGCGGAGCGTTAAACAGAGTG 1017  
| | | | |  
QY 1018 CTGATGCTGTGCTGACGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1077  
| | | | |  
Db 1018 CTGATGCTGTGCTGACGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1077  
| | | | |  
QY 1078 ACCCTTGCATTAATTCATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
| | | | |  
Db 1078 ACCCTTGCATTAATTCATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
| | | | |  
QY 1138 CCGAGAGAGTGGGGATGACGTCAGTCCATGAGCCCTTAAGGGGATGAGGCTTACACG 1197  
| | | | |  
Db 1138 CCGAGAGAGTGGGGATGACGTCAGTCCATGAGCCCTTAAGGGGATGAGGCTTACACG 1197  
| | | | |  
QY 1198 TAAATCAATGCGGCTGACAGAGGGTTCACACCGGAGGGGAGCTAATCTCGAAAGC 1257  
| | | | |  
Db 1198 TAAATCAATGCGGCTGACAGAGGGTTCACACCGGAGGGGAGCTAATCTCGAAAGC 1257  
| | | | |  
QY 1258 GCGTGTATGTCGAGTCCGATGCAAGTCTGCACTGCTGCTGAGAGTCCGAAATGCTAGTAT 1317  
| | | | |  
Db 1258 GCGTGTATGTCGAGTCCGATGCAAGTCTGCACTGCTGCTGAGAGTCCGAAATGCTAGTAT 1317  
| | | | |  
QY 1318 CGCGGATCAGATTCGCGGTGAATTAAGTTCCCGGGCTTGTATACACCGCGCTGACAC 1377  
| | | | |  
Db 1318 CGCGGATCAGATTCGCGGTGAATTAAGTTCCCGGGCTTGTATACACCGCGCTGACAC 1377  
| | | | |  
QY 1378 CATGGAGTGGGTTTCCAGAAAGCAGTATGTTAAACCGTAAAGAGGGGCGCTTCCACG 1437  
| | | | |  
Db 1378 CATGGAGTGGGTTTCCAGAAAGCAGTATGTTAAACCGTAAAGAGGGGCGCTTCCACG 1437  
| | | | |  
QY 1438 TGAGATTCACTGAGGCTG 1457  
| | | | |  
Db 1438 TGAGATTCACTGAGGCTG 1457  
| | | | |

Mon May 8 11:07:21 2006

us-10-659-980a-2.rge

Page 68

Search completed: May 6, 2006, 19:44:56  
Job time : 7523 secs

---



[illegible]

Db	421	AAAGGTTACGGTAAATTAATCGTGA	CCCATGACGGTATGCA	CAGAAAGACACCGCGTAA	C	180					
Qy	481	TACGTCGACAGCCGCGGTAA	TACGTAGAGGTGCA	AGCCTTAATCGAAATTA	CTGGCGCT	540					
Db	481	TACGTGCACAGCCGCGGTAA	TACGTAGAGGTGCA	AGCCTTAATCGAAATTA	CTGGCGCT	540					
Qy	541	AAAGGTCGCGAGGGCGCTT	GTAAAGTCA	AGTGTAAAT	TCCCGGGCTTAA	CTTGGGAAT	600				
Db	541	AAAGGGTCGCGAGGGCGCTT	GTAAAGTCA	AGTGTAAAT	TCCCGGGCTTAA	CTTGGGAAT	600				
Qy	601	TGCGTTTGAACCTAACAAAGCT	AGAGTGTGCGAGAGGAGT	GGAGTGGAAAT	TCCATGCTGTAGCAG		660				
Db	601	TGCGTTTGAACCTAACAAAGCT	AGAGTGTGCGAGAGGAGT	GGAAATTCATGTGTAGCAG		660					
Qy	661	TGAAATGCGTAGAGATAT	TGGAAGAA	CATCGATGCGAAGG	CAGCCTCTG	GGTTAA	CACT	720			
Db	661	TGAAATGCGTAGAGATAT	TGGAAGAA	CATCGATGCGAAGG	CAGCCTCTG	GGTTAA	CACT	720			
Qy	721	GACGCTCATGCA	CGAAAGCGTGGGAG	CAACA	CAGATTA	AGATACCTTGGT	AGTCCA	CGCC	780		
Db	721	GACGCTCATGCA	CGAAAGCGTGGGAG	CAACA	CAGATTA	AGATACCTTGGT	AGTCCA	CGCC	780		
Qy	781	CTAAACGATGTCCAA	CTAGTGTGTGGGCTT	ATTAGGCTTGGTAA	CGAAGCTTAA	CGCGGTGA		840			
Db	781	CTAAACGATGTCCAA	CTAGTGTGTGGGCTT	ATTAGGCTTGGTAA	CGAAGCTTAA	CGCGGTGA		840			
Qy	841	AGTTGACCGCCTG	GGGAGATACGGT	CGCAAGATT	TTAAACTCA	AAAGAA	TTGACG	GGGAG	900		
Db	841	AGTTGACCGCCTG	GGGAGATACGGT	CGCAAGATT	TTAAACTCA	AAAGAA	TTGACG	GGGAG	900		
Qy	901	GCA	CAGCGGTGGAATTA	TGTGGATTAAT	TTCGATGCA	AGCGCA	AAAAA	CCTTAC	CTTAC	960	
Db	901	GCA	CAGCGGTGGAATTA	TGTGGATTAAT	TTCGATGCA	AGCGCA	AAAAA	CCTTAC	CTTAC	960	
Qy	961	GACATGTAGCGAA	TTTCTTAGAGATAG	ATTGTGCTTGGGAG	CGCTTAA	CACAGGTGCTG				1020	
Db	961	GACATGTAGCGAA	TTTCTTAGAGATAG	ATTGTGCTTGGGAG	CGCTTAA	CACAGGTGCTG				1020	
Qy	1021	CATGCGTGTGCT	CAGCTCGTGTGCG	AGATGTTGGGTT	TAAGTCCCGCA	ACGACGCA	ACC			1080	
Db	1021	CATGCGTGTGCT	CAGCTCGTGTGCG	AGATGTTGGGTT	TAAGTCCCGCA	ACGACGCA	ACC			1080	
Qy	1081	CTTGTGATTTA	TTGCTCATCATTTT	GGTGGGCA	CTTTAAT	TGAGACTG	CGCGTGA	CAAA	CCG	1140	
Db	1081	CTTGTGATTTA	TTGCTCATCATTTT	GGTGGGCA	CTTTAAT	TGAGACTG	CGCGTGA	CAAA	CCG	1140	
Qy	1141	GAGAAAGGTGGGAG	TGACGTCA	AGTCCAT	TAGGCCCTTA	TGGGTGAGGCTT	CACACGTAA			1200	
Db	1141	GAGAAAGGTGGGAG	TGACGTCA	AGTCCAT	TAGGCCCTTA	TGGGTGAGGCTT	CACACGTAA			1200	
Qy	1201	TACAATGCGCGT	ACAGAGGGTGTG	CCAACCGCG	CGAGGGGAG	AGCTAAT	CTCAG	AAAAAG	CGCG	1260	
Db	1201	TACAATGCGCGT	ACAGAGGGTGTG	CCAACCGCG	CGAGGGGAG	AGCTAAT	CTCAG	AAAAAG	CGCG	1260	
Qy	1261	TGTAAGTCCGAT	TGCGAGTCTG	CAACTGCA	CTCCGTGA	AGTTCG	GAATCG	CTGTAG	TATGCG	1320	
Db	1261	TGTAAGTCCGAT	TGCGAGTCTG	CAACTGCA	CTCCGTGA	AGTTCG	GAATCG	CTGTAG	TATGCG	1320	
Qy	1321	GGATTCAGAT	TGTCCGCGTGA	ATACGTTCCCGGGT	TTTGTAC	CACCCGCG	CAACAT			1380	
Db	1321	GGATTCAGAT	TGTCCGCGTGA	ATACGTTCCCGGGT	TTTGTAC	CACCCGCG	CAACAT			1380	
Qy	1381	GGAGATGGGTTT	CACAGAA	CAGGTAGTCT	TAA	CCGTAA	GAGAGGGCG	CTTG	GCACG	GTGA	1440
Db	1381	GGAGATGGGTTT	CACAGAA	CAGGTAGTCT	TAA	CCGTAA	GAGAGGGCG	CTTG	GCACG	GTGA	1440
Qy	1441	GATT	CATGACTGGG	GTG	1457						
Db	1441	GATT	CATGACTGGG	GTG	1457						

ADM32705  
ID ADM32705 standard; DNA; 1457 BP.  
XX  
AC ADM32705;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE AOB Type A1 R7clone187 16S rDNA.  
XX  
KW 16S rDNA, ammonia oxidising bacteria; AOB; ammonia; nitrite;  
KW aqueous environment; freshwater; seawater; aquarium; ss.  
XX  
OS Microsomonas sp.  
XX  
PN WO2004026772-A2.  
XX  
PD 01-APR-2004.  
XX  
PF 10-SEP-2003; 2003MO-US028210.  
XX  
PR 19-SEP-2002; 2002US-0386217P.  
PR 19-SEP-2002; 2002US-0386218P.  
PR 19-SEP-2002; 2002US-0386219P.  
XX  
PA (AQUA-) AQUARIA INC.  
XX  
PI Hovanec TA;  
XX  
DR WPI; 2004-304936/28.  
XX  
PT New composition comprising an isolated bacterial strain that oxidizes  
PT ammonia to nitrite, useful for alleviating or preventing the accumulation  
PT of ammonia in aqueous environment.  
XX  
PS Disclosure; Page 10; 98pp; English.  
XX  
XX This sequence represents a 16S rDNA sequence derived from an ammonia  
CC oxidising bacteria (AOB). This sequence may be used in a composition  
CC which comprises an isolated bacterial strain that oxidises ammonia to  
CC nitrite. The composition may be used for alleviating or preventing the  
CC accumulation of ammonia in a medium. The ammonia is reduced by at least  
CC 30% when compared with a level of ammonia that would exist in the absence  
CC of the bacterial strain. The composition is useful for alleviating or  
CC preventing the accumulation of ammonia in aqueous environment, e.g. a  
CC freshwater or seawater aquarium.  
XX  
SQ Sequence 1457 BP; 379 A; 317 C; 455 G; 306 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1457; DB 12; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATTTGAACGCTGGCGCATGCTTTTACACATGCAAGTCGAAACGGACGACGATGCTTGAT 60  
DB 1 ATTTGAACGCTGGCGCATGCTTTTACACATGCAAGTCGAAACGGACGATGCTTGAT 60  
QY 61 CTGTGTGCGAATGCGGACGGGTGAGTAATGCAATCGAAACGTATCCGAAAGAGGGGGTA 120  
DB 61 CTGTGTGCGAATGCGGACGGGTGAGTAATGCAATCGAAACGTATCCGAAAGAGGGGGTA 120  
QY 121 ACGCATGAAAGATGTGCTAATACCGCATATACTCTAAGAGAAAGCGAGATCGAAA 180  
DB 121 ACGCATGAAAGATGTGCTAATACCGCATATACTCTAAGAGAAAGCGAGATCGAAA 180  
QY 181 GACCTTGGCTTTTGGACGGCCGATGTCTGATTAGCTAAGTTGGTGGGTAAAGGCTTAC 240  
DB 181 GACCTTGGCTTTTGGACGGCCGATGTCTGATTAGCTAAGTTGGTGGGTAAAGGCTTAC 240  
QY 241 CAAGCGGACGATCACTAGTTGGTCTGAGAGAGACGACCGACCTGGGACCTGAGACACG 300  
DB 241 CAAGCGGACGATCACTAGTTGGTCTGAGAGAGACGACCGACCTGGGACCTGAGACACG 300  
QY 301 GCCGAGACTCTTACGGAGAGCGACGATGGGGAAATTTTGGACATAGGCGGCACGCTGATC 360

DB 301 GCCGAGACTCTTACGGAGAGCGACGATGGGGAAATTTTGGACATAGGCGGCACGCTGATC 360  
QY 361 CAGCAATGCGCGGTGAGTAGAAGAGCCCTTGCGGGTTTGAAGCTCTTTCACTGAGAGAA 420  
DB 361 CAGCAATGCGCGGTGAGTAGAAGAGCCCTTGCGGGTTTGAAGCTCTTTCACTGAGAGAA 420  
QY 421 AAAGTTACGTTAATATCTGACCCATGACGGTATCGACGAAGAAGCACCGGCTAAC 480  
DB 421 AAAGTTACGTTAATATCTGACCCATGACGGTATCGACGAAGAAGCACCGGCTAAC 480  
QY 481 TAGGTCCAGACGCGCGGTAAATATGAGGTGACGACCTTATCGGAATTACTGGGCGT 540  
DB 481 TAGGTCCAGACGCGCGGTAAATATGAGGTGACGACCTTATCGGAATTACTGGGCGT 540  
QY 541 AAAGGTGCGCAGCGCGCCCTTGTAACTCAGATGTGAATCCCGGGCTTAACTCGGAAAT 600  
DB 541 AAAGGTGCGCAGCGCGCCCTTGTAACTCAGATGTGAATCCCGGGCTTAACTCGGAAAT 600  
QY 601 TCGCTTTGAACTTACAAAGCTAGAGTGTGGCAGAGGAGGTGAAATTCATGTGTAGCAG 660  
DB 601 TCGCTTTGAACTTACAAAGCTAGAGTGTGGCAGAGGAGGTGAAATTCATGTGTAGCAG 660  
QY 661 TGAATGCGTAGATATGGAAGAAATCATGATGCGAAGGACGCTCTGGGTTAACACT 720  
DB 661 TGAATGCGTAGATATGGAAGAAATCATGATGCGAAGGACGCTCTGGGTTAACACT 720  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCCAGCC 780  
DB 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCCAGCC 780  
QY 781 CTTAAACGATGTCAATGATTGTTGGGCTTATTAAGCTTGTGTAACGAAAGTAAACGGTGA 840  
DB 781 CTTAAACGATGTCAATGATTGTTGGGCTTATTAAGCTTGTGTAACGAAAGTAAACGGTGA 840  
QY 841 AGTTGACCGCTTGGGAGATACCGTCCGAAATTAATCAAGAAATTGACGGGAGACC 900  
DB 841 AGTTGACCGCTTGGGAGATACCGTCCGAAATTAATCAAGAAATTGACGGGAGACC 900  
QY 901 GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGGGAAACCTTACCTACCTT 960  
DB 901 GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGGGAAACCTTACCTACCTT 960  
QY 961 GACATGTAGGGAATTTTCTAGAGATAGATTAGTTCGGGAAACGCTTAAACAGGTGCTG 1020  
DB 961 GACATGTAGGGAATTTTCTAGAGATAGATTAGTTCGGGAAACGCTTAAACAGGTGCTG 1020  
QY 1021 CATGCTGTGTGACGCTCGTGTGATGATGTTGGGTTAAGTCCCGCAACGACGCAACC 1080  
DB 1021 CATGCTGTGTGACGCTCGTGTGATGATGTTGGGTTAAGTCCCGCAACGACGCAACC 1080  
QY 1081 CTTGTCAATTAATTCGATCACTATTGTTGGGCACTTTAATGAACTGCGGTGACAAACCG 1140  
DB 1081 CTTGTCAATTAATTCGATCACTATTGTTGGGCACTTTAATGAACTGCGGTGACAAACCG 1140  
QY 1141 GAGGAGGTGGGAGTGAAGTCAAGTCTCATGCGCTTATAGGGGTTCACAGCTAA 1200  
DB 1141 GAGGAGGTGGGAGTGAAGTCAAGTCTCATGCGCTTATAGGGGTTCACAGCTAA 1200  
QY 1201 TACAAATGCGCGCTTACAGAGGTTTCCCAACCCCGAGGGGAGGATCTAGAAAGGCGC 1260  
DB 1201 TACAAATGCGCGCTTACAGAGGTTTCCCAACCCCGAGGGGAGGATCTAGAAAGGCGC 1260  
QY 1261 TCGTAGTCCGAGATCGAGTCTGCAACTCGAATCGGTAAGTCCGTAATCGC 1320  
DB 1261 TCGTAGTCCGAGATCGAGTCTGCAACTCGAATCGGTAAGTCCGTAATCGC 1320  
QY 1321 GGATCAGCATGTGCGGTGTAATAGTTCCCGGGCTTTGTAACACCGCCGTCACACCAT 1380  
DB 1321 GGATCAGCATGTGCGGTGTAATAGTTCCCGGGCTTTGTAACACCGCCGTCACACCAT 1380  
QY 1381 GGGAGTGGGTTTCCAGGAAGCAGTGTGTTAACCGTTAAGAGGGCGCTTGCCAGGCTGA 1440

Db 1381 GGGAGTGGGTTTCACCAAGACGATAGTCTAAACGTAAAGAGGGGCTTGCCACGGTGA 1440  
QY 1441 GATTCACTGAGGGTG 1457  
Db 1441 GATTCACTGAGGGTG 1457  
RESULT 3  
ABA02416  
ID ABA02416 standard; DNA; 1457 BP.  
XX  
XX ABA02416;  
AC  
XX 29-AUG-2003 (revised)  
DT 04-MAR-2002 (first entry)  
XX  
XX Type A ammonia-oxidising bacterium 16S rRNA gene sequence, R7clone140.  
DE  
XX Type A; ammonia-oxidising bacterium; AOB; nitrite; 16S rRNA gene;  
KM ribosomal RNA; aquarium; aquaculture; waste water treatment;  
KW bioremediation; ds.  
XX  
XX Nitrosomonadales.  
OS  
XX MO200190312-A1.  
PN  
XX 29-NOV-2001.  
PD  
XX 17-MAY-2001; 2001MO-US016265.  
PF  
XX 19-MAY-2000; 2000US-00573684.  
PR  
XX (AQUA-) AQUARIA INC.  
PA  
XX Hovanec TA, Burrell PC;  
PI  
XX MPI; 2002-075367/10.  
DR  
XX  
XX New bacteria capable of oxidizing ammonia to nitrite, for preventing or  
PT alleviating the accumulation of ammonia in fresh water aquaria, seawater  
PT aquaria and waste water.  
XX  
XX  
PS Claim 2; Page 5; 62pp; English.  
CC The invention relates to 4 novel types of ammonia-oxidising bacteria  
CC (AOB) found in freshwater aquaria. The bacteria are able to oxidise  
CC ammonia to nitrite and are members of the ammonia-oxidising bacteria  
CC family of the beta subdivision of Proteobacteria. The 4 types of bacteria  
CC can be distinguished on the basis of their 16S rRNA (ribosomal RNA) gene  
CC sequences (ABA02416-ABA02419), and are classified as AOB type A (e.g.,  
CC R7clone140), type A1 (e.g., R7clone187), type B (e.g., R7clone5) and type  
CC C (e.g., R7clone47). The invention also encompasses isolated 16S rRNA  
CC gene sequences of the ammonia-oxidising bacteria of the invention,  
CC oligonucleotide probes and primers for the detection of these bacteria,  
CC and compositions comprising the bacteria. The bacteria of the invention  
CC are useful in biological filters for reducing ammonia accumulation in  
CC both freshwater and seawater aquaria. They may also be used in waste  
CC water treatment and in bioremediation processes to reduce the level of  
CC pollution caused by ammonia. The present sequences represent R7clone140,  
CC a 16S rRNA gene sequence from the type A ammonia-oxidising bacterium of  
CC the invention. (Updated on 29-AUG-2003 to standardise OS field)  
XX  
XX  
SQ Sequence 1457 BP; 378 A; 315 C; 456 G; 308 T; 0 U; 0 Other:  
Query Match 99.7%; Score 1452.2; DB 6; Length 1457;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTGAACGGCAGACAGGATGCTTGCAAT 60  
Db 1 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTGAACGGCAGACAGGATGCTTGCAAT 60  
QY 61 CTGGTGGCGAGTGGCGGAGCGGTGATGATGATCGGAACGTATCCAGAAAGAGGGGGGTA 120

Db 61 CTGGTGGCGAGTGGCGGAGCGGTGATGATGATGATCGGAACGTATCCAGAAAGAGGGGGGTA 120  
QY 121 ACGCATCGAAAGATGTGCTAATAATCCGCATATATCTTAAAGAGAAAGCAGGGGATCGAAA 180  
Db 121 ACGCATCGAAAGATGTGCTAATAATCCGCATATATCTTAAAGAGAAAGCAGGGGATCGAAA 180  
QY 181 GACCTTGGCCTTTTGGAGCGCGCATATGCTGATTTAGCTTGTGTGGGGTAAAGGCTTAC 240  
Db 181 GACCTTGGCCTTTTGGAGCGCGCATATGCTGATTTAGCTTGTGTGGGGTAAAGGCTTAC 240  
QY 241 CAAGCGACGATCAGTATGTTGCTTGAAGAGACGACACGACCACTGGGACTGAGACACG 300  
Db 241 CAAGCGACGATCAGTATGTTGCTTGAAGAGACGACACGACCACTGGGACTGAGACACG 300  
QY 301 GCCCAGACTCTTACGGGAGCAGCAGTGGGGAAATTTTGGACAAATGGCGCCAAAGCTGATC 360  
Db 301 GCCCAGACTCTTACGGGAGCAGCAGTGGGGAAATTTTGGACAAATGGCGCCAAAGCTGATC 360  
QY 361 CAGCAATGCGGGTGTGATGAAGAGGCTTGGGGTGTTAAAGCTCTTTCAGTCGAGAAAG 420  
Db 361 CAGCAATGCGGGTGTGATGAAGAGGCTTGGGGTGTTAAAGCTCTTTCAGTCGAGAAAG 420  
QY 421 AAAGGTTACGTTAAATATCTGACCCATGACGATATCGACAAAGAAAGACCGGCTAAC 480  
Db 421 AAAGGTTACGTTAAATATCTGACCCATGACGATATCGACAAAGAAAGACCGGCTAAC 480  
QY 481 TACGTGCCAGACCGCGGCTAATACGTAGGGTGTGCAAGGCTTAAATGGAATTACTGGGCGT 540  
Db 481 TACGTGCCAGACCGCGGCTAATACGTAGGGTGTGCAAGGCTTAAATGGAATTACTGGGCGT 540  
QY 541 AAAGGGTGCAGGGGCGCTTGTAAAGTGAATGCAATGTAATCCCGGGCTTAACTGAGAAAT 600  
Db 541 AAAGGGTGCAGGGGCGCTTGTAAAGTGAATGCAATGTAATCCCGGGCTTAACTGAGAAAT 600  
QY 601 TCGGTTTGAATCAAAAGCTAGAGTGTGCGCAGAGGAGTGAATTCATGCTGTAGCAG 660  
Db 601 TCGGTTTGAATCAAAAGCTAGAGTGTGCGCAGAGGAGTGAATTCATGCTGTAGCAG 660  
QY 661 TGAATTCGTAGATATGGAAGAAATCATGATGCGGAAGCGCCTCTGGGTTAACT 720  
Db 661 TGAATTCGTAGATATGGAAGAAATCATGATGCGGAAGCGCCTCTGGGTTAACT 720  
QY 721 GACGCTCATGCAAGAAAGGTGGGAGCAAGCAAGGATTAATACCTGCTAGTCCACGCC 780  
Db 721 GACGCTCATGCAAGAAAGGTGGGAGCAAGCAAGGATTAATACCTGCTAGTCCACGCC 780  
QY 781 CTAAACGATGCTAATGTTGTTGGGCTTATTAGCTTGTGTTGAAGCTTAACCGCTGA 840  
Db 781 CTAAACGATGCTAATGTTGTTGGGCTTATTAGCTTGTGTTGAAGCTTAACCGCTGA 840  
QY 841 AGTTGACCGCTCGGGAGATACGCTGCGCAAGATTAATACTCAAGAAATTTGACGGGAGCC 900  
Db 841 AGTTGACCGCTCGGGAGATACGCTGCGCAAGATTAATACTCAAGAAATTTGACGGGAGCC 900  
QY 901 GACAAGGGGTGATTAATGATGATTAATTCATGTCAGACGCGGAAACCTTACTCCCTT 960  
Db 901 GACAAGGGGTGATTAATGATGATTAATTCATGTCAGACGCGGAAACCTTACTCCCTT 960  
QY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAATGCTTGGGAAACGCTAACACAGGTGCTG 1020  
Db 961 GACATGTAGCGAATTTTCTAGAGATAGATTAATGCTTGGGAAACGCTAACACAGGTGCTG 1020  
QY 1021 CATGCTGTGCTGACGCTGTGCTGAGATGTTGGGTTAATGTCGCGCAACGAGCCCAAC 1080  
Db 1021 CATGCTGTGCTGACGCTGTGCTGAGATGTTGGGTTAATGTCGCGCAACGAGCCCAAC 1080  
QY 1081 CTGTGCTTAATATGCAATATTTGTTGGGCACTTAAATGAGACTGGCGGTGACAAACCG 1140  
Db 1081 CTGTGCTTAATATGCAATATTTGTTGGGCACTTAAATGAGACTGGCGGTGACAAACCG 1140  
QY 1141 GAGAAAGGTGGGATGAGCTCAAGTCTCATGCGCCCTTATGGGTTAGGGCTTACACGTA 1200

Db 1141 GAGGAAGTGGGATGATGCTCAAGTCTCTATGCGCCCTTATGGTAGGGCTTTCACAGCTAA 1200  
QY 1201 TACAAATGGCGCGCTACAGAGGGTTGCCAACCCGGAGGGGAGCTAATCTCAGAAAGCGC 1260  
Db 1201 TACAAATGGCGCGCTACAGAGGGTTGCCAACCCGGAGGGGAGCTAATCTCAGAAAGCGC 1260  
QY 1261 TCGTAGTCCGGATCGAGTCTGCAACTGCACTCCGCTGAAGTCCGAATCGCTAGTAATCGC 1320  
Db 1261 TCGTAGTCCGGATCGAGTCTGCAACTGCACTCCGCTGAAGTCCGAATCGCTAGTAATCGC 1320  
QY 1321 GGAATCAGCATGTGCGCGGTGAATACGTTCCCGGGCTTGTACACACCGCCGCTCAGACAT 1380  
Db 1321 GGAATCAGCATGTGCGCGGTGAATACGTTCCCGGGCTTGTACACACCGCCGCTCAGACAT 1380  
QY 1381 GGGAGTGGGTTTTCACGAAAGCAGGTAGTCTAACCCGTAAGAGAGGCGCTTGCCACGCTGA 1440  
Db 1381 GGGAGTGGGTTTTCACGAAAGCAGGTAGTCTAACCCGTAAGAGAGGCGCTTGCCACGCTGA 1440  
QY 1441 GATTTCATGACTGGGGTG 1457  
Db 1441 GATTTCATGACTGGGGTG 1457

RESULT 4  
ADM32704  
ID ADM32704 standard; DNA; 1457 BP.

XX ADM32704;

XX 17-JUN-2004 (first entry)

XX AOB Type A R7clone140 16S rDNA.

XX 16S rDNA; ammonia oxidising bacteria; AOB; ammonia; nitrite;

KM aqueous environment; freshwater; seawater; aquarium; ss.

XX Nitrosomonas sp.

PN W02004026772-A2.

XX 01-APR-2004.

XX 10-SEP-2003; 2003WO-US028210.

XX 19-SEP-2002; 2002US-0386217P.

PR 19-SEP-2002; 2002US-0386218P.

PR 19-SEP-2002; 2002US-0386219P.

XX (AQUA-) AQUARIA INC.

XX Hovaneec TA;

DR WPI; 2004-304936/28.

XX New composition comprising an isolated bacterial strain that oxidizes

PT ammonia to nitrite, useful for alleviating or preventing the accumulation

XX of ammonia in aqueous environment.

PS Disclosure; Page 8-9; 98pp; English.

XX This sequence represents a 16S rDNA sequence derived from an ammonia

CC oxidizing bacteria (AOB). This sequence may be used in a composition

CC which comprises an isolated bacterial strain that oxidizes ammonia to

CC nitrite. The composition may be used for alleviating or preventing the

CC accumulation of ammonia in a medium. The ammonia is reduced by at least

CC 30% when compared with a level of ammonia that would exist in the absence

CC of the bacterial strain. The composition is useful for alleviating or

CC preventing the accumulation of ammonia in aqueous environment, e.g. a

CC freshwater or seawater aquarium.

QY Sequence 1457 BP; 378 A; 315 C; 456 G; 308 T; 0 U; 0 Other;

QY Query Match 99.7%; Score 1452.2; DB 12; Length 1457;

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAACGCTGGCGCATCTTTTACCATGCAAGTGAACCGGACAGACCGGATCGTCAAT 60  
Db 1 ATGGAACGCTGGCGCATCTTTTACCATGCAAGTGAACCGGACAGACCGGATCGTCAAT 60  
QY 61 CTGGTGGCGAGTGGCGGAGGAGTGAATGCAATCGGAACGTAATCGGAAGAGGGGGTGA 120  
Db 61 CTGGTGGCGAGTGGCGGAGGAGTGAATGCAATCGGAACGTAATCGGAAGAGGGGGTGA 120  
QY 121 ACCGATCGAAAGATGTGCTAATACCGCATATCTCTTAAGAGGAAAGCAGGGGATCGAAA 180  
Db 121 ACCGATCGAAAGATGTGCTAATACCGCATATCTCTTAAGAGGAAAGCAGGGGATCGAAA 180  
QY 181 GACCTTGCGCTTTTGGAGCGGCGGATGCTGATTAAGTCTAGTGGGTGAAGGCTTAC 240  
Db 181 GACCTTGCGCTTTTGGAGCGGCGGATGCTGATTAAGTCTAGTGGGTGAAGGCTTAC 240  
QY 241 CAAGGCGACGATCAATGATGCTGAGAGAGACACGACACTGGGAGCTGAGACAG 300  
Db 241 CAAGGCGACGATCAATGATGCTGAGAGAGACACGACACTGGGAGCTGAGACAG 300  
QY 301 GCCAGACCTCTACGCGGAGGACACAGTGGGAAATTTTGGACAAATGGCGCAGGCTGATC 360  
Db 301 GCCAGACCTCTACGCGGAGGACACAGTGGGAAATTTTGGACAAATGGCGCAGGCTGATC 360  
QY 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTTTTCAGTGAAGAGA 420  
Db 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTTTTCAGTGAAGAGA 420  
QY 421 AAAGTTACGCTTAATTAATGTCACCCATGACGTAACGTAACGAGAAAGACCGGCTTAC 480  
Db 421 AAAGTTACGCTTAATTAATGTCACCCATGACGTAACGTAACGAGAAAGACCGGCTTAC 480  
QY 481 TAAGTGCACAGCGCGGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 540  
Db 481 TAAGTGCACAGCGCGGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 540  
QY 541 AAAGGATGCGCAGGCGGCTTGTAAAGTGAATGTAATGTAATGTAATGTAATGTAATG 600  
Db 541 AAAGGATGCGCAGGCGGCTTGTAAAGTGAATGTAATGTAATGTAATGTAATGTAATG 600  
QY 601 TGGCTTTGAAACTTAAGCTTAAGTGGTGGCAGAGGAGGTAATCCATGTTAGCAG 660  
Db 601 TGGCTTTGAAACTTAAGCTTAAGTGGTGGCAGAGGAGGTAATCCATGTTAGCAG 660  
QY 661 TGAATGCGTAGAGATATGGAAGAACATGATGCGAAGGACGCTCTGGGTTAACACT 720  
Db 661 TGAATGCGTAGAGATATGGAAGAACATGATGCGAAGGACGCTCTGGGTTAACACT 720  
QY 721 GACGCTCATGACCAAAAGCGTGGGAGCAAAACAGATTAAGTACCTTGTAATGTCAGCC 780  
Db 721 GACGCTCATGACCAAAAGCGTGGGAGCAAAACAGATTAAGTACCTTGTAATGTCAGCC 780  
QY 781 CTAAACGATGCAACTGATGTTGGGCTTAATTAAGCTTGGTAAAGCAAGCTAACGCGGA 840  
Db 781 CTAAACGATGCAACTGATGTTGGGCTTAATTAAGCTTGGTAAAGCAAGCTAACGCGGA 840  
QY 841 AGTTGACCGCTGGGGAGTACGTCGCAAGTAAACTCAAAAGAAATGACGGGAGCC 900  
Db 841 AGTTGACCGCTGGGGAGTACGTCGCAAGTAAACTCAAAAGAAATGACGGGAGCC 900  
QY 901 GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAATAAAGCTTACCTT 960  
Db 901 GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAATAAAGCTTACCTT 960  
QY 961 GACATGTAAGCAATTTTCTAAGATTAAGTATGATGATGATGATGATGATGATGATG 1020  
Db 961 GACATGTAAGCAATTTTCTAAGATTAAGTATGATGATGATGATGATGATGATGATG 1020  
QY 1021 CATGGCTGCTGACGCTGCTGCTGCTGATGATGTTGGGTTAAGTCCCGCAACGAGCGAAC 1080  
Db 1021 CATGGCTGCTGACGCTGCTGCTGCTGATGATGTTGGGTTAAGTCCCGCAACGAGCGAAC 1080

Db 1021 CATGCTGTCGTACGCTCGTGTGTGATGTTGGTTAATGCCGGAACGAGCGAACC 1080  
Qy 1081 CTGTGATTAATATGTCATCTTTGGTGGACCTTTAATAGACTGCCGGTGAACAAACCG 1140  
Db 1081 CTGTGATTAATATGTCATCTTTGGTGGACCTTTAATAGACTGCCGGTGAACAAACCG 1140  
Qy 1141 GAGGAGGTGGGATGATGACGTCAAGTCTCTCATGGCCCTTAATGAGGCTTTCACAGCTAA 1200  
Db 1141 GAGGAGGTGGGATGATGACGTCAAGTCTCTCATGGCCCTTAATGAGGCTTTCACAGCTAA 1200  
Qy 1201 TACAAATGGCGCGTACAGAGGGTTCACAAACCGGAGGGAGCTTAATTCAGAAAGCGCG 1260  
Db 1201 TACAAATGGCGCGTACAGAGGGTTCACAAACCGGAGGGAGCTTAATTCAGAAAGCGCG 1260  
Qy 1261 TCGTATGTCGGATTCGAGATTCGCACTCGACTCCGTAAGTCCGAATCCGTATTAATCCG 1320  
Db 1261 TCGTATGTCGGATTCGAGATTCGCACTCGACTCCGTAAGTCCGAATCCGTATTAATCCG 1320  
Qy 1321 GGATCAGCATGTGCGCGGTGAATAGTTCCGCGGTCTTGTACACACGCCCGTCAACCAT 1380  
Db 1321 GGATCAGCATGTGCGCGGTGAATAGTTCCGCGGTCTTGTACACACGCCCGTCAACCAT 1380  
Qy 1381 GGGAGTGGGTTTACAGAAAGCAGTAAGTCTAAACGTAAGAGAGGCGCTTTCACAGCTGA 1440  
Db 1381 GGGAGTGGGTTTACAGAAAGCAGTAAGTCTAAACGTAAGAGAGGCGCTTTCACAGCTGA 1440  
Qy 1441 GATTCAATGACTGGGGTG 1457  
Db 1441 GATTCAATGACTGGGGTG 1457

RESULT 5  
ADM32723  
ID ADM32723 standard; DNA; 1491 BP.

XX ADM32723;  
XX 17-JUN-2004 (first entry)  
XX AOB Bf1c1one57 16S rDNA.  
XX 16S rDNA; ammonia oxidising bacteria; AOB; ammonia; nitrite;  
XX aqueous environment; freshwater; seawater; aquarium; ss.  
XX Nitrosomonas aestuarii.  
XX WO2004026772-A2.  
XX 01-APR-2004.  
XX 10-SEP-2003; 2003WO-US028210.  
XX 19-SEP-2002; 2002US-0386217P.  
XX 19-SEP-2002; 2002US-0386218P.  
XX 19-SEP-2002; 2002US-0386219P.  
XX (AQUA-) AQUARIA INC.  
XX Hovanez TA;  
XX MPI; 2004-304936/28.  
XX New composition comprising an isolated bacterial strain that oxidizes  
XX ammonia to nitrite, useful for alleviating or preventing the accumulation  
XX of ammonia in aqueous environment.  
XX Claim 1; Page 15; 98pp; English.  
XX This sequence represents a 16S rDNA sequence derived from an ammonia  
XX oxidising bacteria (AOB). This sequence may be used in a composition  
XX which comprises an isolated bacterial strain that oxidizes ammonia to  
XX nitrite. The composition may be used for alleviating or preventing the  
XX accumulation of ammonia in a medium. The ammonia is reduced by at least

CC 30% when compared with a level of ammonia that would exist in the absence  
CC of the bacterial strain. The composition is useful for alleviating or  
CC preventing the accumulation of ammonia in aqueous environment, e.g. a  
CC freshwater or seawater aquarium.

XX Sequence 1491 BP; 392 A; 323 C; 457 G; 319 T; 0 U; 0 Other;

Query Match 95.4%; Score 1390; DB 12; Length 1491;  
Best Local Similarity 97.5%; Pred. No. 0;  
Matches 1422; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

Qy 1 ATTGAACGTGCGCGGATGCTTTAACAATGACAACTGAAACGCGACGCGATGCTTGAT 60  
Db 18 ATTGAACGTGCGCGGATGCTTTAACAATGACAACTGAAACGCGACGCGATGCTTGAT 77  
Qy 61 CTGTGGCAGATGGCGGACGCGGTGATTAATGATGAAAGCTTAATCAAGAGGGGGGTA 120  
Db 78 CTGTGGCAGATGGCGGACGCGGTGATTAATGATGAAAGCTTAATCAAGAGGGGGGTA 137  
Qy 121 AGGCATCGAAAGATGCTTAATACGCATTAATCTTAAGAGGAAAGAGGGGATCGAA 180  
Db 138 AGGCATCGAAAGATGCTTAATACGCATTAATCTTAAGAGGAAAGAGGGGATCGAA 197  
Qy 181 GACCTTGCCTTTTGAAGCGCCGATGCTGATTAAGCTAATGTTGGGTAAAGCTTAC 240  
Db 198 GACCTTGCCTTTTGAAGCGCCGATGCTGATTAAGCTAATGTTGGGTAAAGCTTAC 257  
Qy 241 CAAGCGACGATCAGTAATGTTGTTGAGAGAGACAGACAGCCACTGGGACTGAGACAG 300  
Db 258 CAAGCGACGATCAGTAATGTTGTTGAGAGAGACAGACAGCCACTGGGACTGAGACAG 317  
Qy 301 GCCCAGACTCTTACCGGAGGAGCAGATGGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360  
Db 318 GCCCAGACTCTTACCGGAGGAGCAGATGGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 377  
Qy 361 CAGCAATGCGCGGTGATGTAAGAGAGCCCTTCGGGTGTTAAAGCTTTTCAGTCGAGAGA 420  
Db 378 CAGCAATGCGCGGTGATGTAAGAGAGCCCTTCGGGTGTTAAAGCTTTTCAGTCGAGAGA 437  
Qy 421 AAAGGTTACGTAATTAATCGTACCATGACGATGATGCAAGAGAAAGACCCGCTAAC 480  
Db 438 AAAGGTTACGTAATTAATCGTACCATGATGATGCAAGAGAAAGACCCGCTAAC 497  
Qy 481 TAGGTGCCAGACCGCGGTAAATACGTAGGGGTGAAGGCTTAATGCGAATTAATCTGGGCT 540  
Db 498 TAGGTGCCAGACCGCGGTAAATACGTAGGGGTGAAGGCTTAATGCGAATTAATCTGGGCT 557  
Qy 541 AAAGGTGGCAGAGCGGCTTGTAAATGATGTAATTCGCCGGCTTAACTTGGGAAT 600  
Db 558 AAAGGTGGCAGAGCGGCTTGTAAATGATGTAATTCGCCGGCTTAACTTGGGAAT 617  
Qy 601 TCGCTTTGAAACTTAAGAGCTGAGTGGCGAGAGAGTGGGAATTCATGTGAGCAG 660  
Db 618 TCGCTTTGAAACTTAAGAGCTGAGTGGCGAGAGAGTGGGAATTCATGTGAGCAG 677  
Qy 661 TGAATGCGTAGATGATGAGAAACATGATGCGAAGGCGAGCTTCCGTGGTTAACT 720  
Db 678 TGAATGCGTAGATGATGAGAAACATGATGCGAAGGCGAGCTTCCGTGGTTAACT 737  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGAGATTAGATACCTGTAGTCCACGCC 780  
Db 738 GACGCTCATGACGAAAGCGTGGGAGCAACAGAGATTAGATACCTGTAGTCCACGCC 797  
Qy 781 CTAAACGATGTAATGATGTTGGGCTTATTAAGCTTGGTAAGAAAGCTTAACCGCTGA 840  
Db 798 CTAAACGATGTAATGATGTTGGGCTTATTAAGCTTGGTAAGAAAGCTTAACCGCTGA 857  
Qy 841 AGTGACGCGCTGGGAGATGAGGTGCAAGATTAACTCAAGAGATTGACGGGAGCC 900  
Db 858 AGTGACGCGCTGGGAGATGAGGTGCAAGATTAACTCAAGAGATTGACGGGAGCC 917  
Qy 901 GCACAGGCGGTGATTAATGATGATTAATTCGATGCAAGCGGAAACCTTAACCTACCTT 960



Db	918	GCAACAGCGGTGACATTATGTGAGATTAAATTCAGTACGACGCAAAAAACCTTAACCTT	977
Qy	961	GACATGTAGCCAAATTTTCTAGAGATGAAATTAAGT-CTTCGGGAAACGTTAACACAGTGTCT	1019
Db	978	GACATGTATGCAATATATTTTAGAGATAAATAATGTGCTTCGGGAAACGTTAACACAGTGTCT	1037
Qy	1020	GCATGCTGTGTGTCAAGTCTGTGTGTGTGAGATGTTTGGGTTAACTCCCGCAAGACCGCAAC	1079
Db	1038	GCATGCTGTGTGTCAAGTCTGTGTGTGTGAGATGTTTGGGTTAACTCCCGCAAGACCGCAAC	1097
Qy	1080	CCTTGTCATTATATTCGCATATTTGGTTGGGCACTTTAATGAGACTGCGGTGACAAAC	1139
Db	1098	CCTTGTCATTATATTCGCATATTTAGTTGGGCACTTTAATGAGACTGCGGTGACAAAC	1157
Qy	1140	GGAGGAAAGTGGGAGATGACGTCAAGTCTCATAGCGCTTATGGGTAGGGCTTACACGTA	1199
Db	1158	GGAGGAAAGTGGGAGATGACGTCAAGTCTCATAGCGCTTATGGGTAGGGCTTACACGTA	1217
Qy	1200	ATACATATGGCCGCTACAGAGGTTGCCAACCCCGGAGGGGAGACATACTCAGAAAGGCG	1258
Db	1218	ATACATATGGCCGCTACAGAGGTTGCCAACCCCGGAGGGGAGACATACTCAGAAAGGCG	1277
Qy	1260	GTCGTAATCCGGATGGAGTCTGCAACTCGAATCCGTGAAGTCGGAATCGTAGTAATCG	1319
Db	1278	GTCGTAATCCGGATGGAGTCTGCAACTCGAATCCGTGAAGTCGGAATCGTAGTAATCG	1337
Qy	1320	CGGATCAGCATGTCCGCGTGAATACGTTCCCGGATCTTGTACACCGCCCGTCAACCA	1379
Db	1338	CGGATCAGCATGTCCGCGTGAATACGTTCCCGGATCTTGTACACCGCCCGTCAACCA	1397
Qy	1380	TGGAGTGGGTTTCAACGAGAGCAGTACTTAAACCGTAAAGAGGGCGCTTGCCACGGTG	1438
Db	1398	TGGAGTGGGTTTCAACGAGAGCAGTACTTAAACCGTAAAGAGGGCGCTTGCCACGGTG	1456
Qy	1440	AGATTCATGATCGGGGTG 1457	
Db	1458	AGATTCATGATCGGGGTG 1475	
RESULT 6			
ADM32722			
ID	ADM32722	Standard; DNA; 1494 BP.	
AC	ADM32722;		
XX			
XX			
DT	17-JUN-2004	(first entry)	
XX			
DE	AOB P4clone31 16S rDNA.		
XX			
KW	16S rDNA; ammonia oxidising bacteria; AOB; ammonia; nitrite;		
XX	aqueous environment; freshwater; seawater; aquarium; ss.		
OS	Nitrosomonas aestuarii.		
XX			
PN	WO2004026772-A2.		
XX			
PD	01-APR-2004.		
XX			
PF	10-SEP-2003; 2003WO-US028210.		
XX			
PR	19-SEP-2002; 2002US-0386217P.		
XX			
PR	19-SEP-2002; 2002US-0386218P.		
XX			
PR	19-SEP-2002; 2002US-0386219P.		
XX			
PA	(AQUA-) AQUARIA INC.		
XX			
PI	Hovanec TA;		
XX			
DR	WPI; 2004-304936/28.		
XX			
PT	New composition comprising an isolated bacterial strain that oxidizes		
PT	ammonia to nitrite, useful for alleviating or preventing the accumulation		
PT	of ammonia in aqueous environment.		

PS Claim 1; Page 14; 98bp; English.

XX This sequence represents a 16S rDNA sequence derived from an ammonia  
PS oxidizing bacteria (AOB). This sequence may be used in a composition  
XX which comprises an isolated bacterial strain that oxidizes ammonia to  
CC nitrite. The composition may be used for alleviating or preventing the  
CC accumulation of ammonia in a medium. The ammonia is reduced by at least  
CC 30% when compared with a level of ammonia that would exist in the absence  
CC of the bacterial strain. The composition is useful for alleviating or  
CC preventing the accumulation of ammonia in aqueous environment, e.g. a  
CC freshwater or seawater aquarium.

XX  
SQ Sequence 1494 BP; 392 A; 327 C; 459 G; 316 T; 0 U; 0 Other;

Query Match 93.2%; Score 1357.6; DB 12; Length 1494;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 1401; Conservative 0; Mismatches 54; Indels 1; Gaps 1

1 ATTGAACGCTGGCGGCATGCTTTACATGCAATGCAAGCGACACGATGCTTGCAT 60  
19 ATTGAACGCTGGCGGCATGCTTTACATGCAATGCAAGCGACACGATGCTTGCAC 78  
61 CTGTGCGCGAGTGTGCGGACGCGGTGATGATGATGCAACGTTATCAGAAAGCGGAGTA 120  
79 CTGTGCGCGAGTGTGCGGACGCGGTGATGATGATGCAACGTTATCAGAAAGCGGAGTA 138  
121 ACGCATCGAAAGATGTGCTTAATACCGCATATCTTAAAGAGGAAACGAGGGATTCGAAA 180  
139 ACGCATCGAAAGATGTGCTTAATACCGCATATCTTAAAGAGGAAACGAGGGATTCGAAA 198  
181 GACCTTGGCGCTTTTGGAGCGCGCCCATATGCTGATTGAGTGTGTGGGGTTAAAGGCTAC 240  
199 GACCTTGGCGCTTTTGGAGCGCGCCCATATGCTGATTGAGTGTGTGGGGTTAAAGGCTTAC 258  
241 CAAGGCGACATCAGATGTGTGTCTGAGAGAGACGACACGACCACTGGGACTGAGACAG 300  
259 CAAGGCGACATCAGATGTGTGTCTGAGAGAGACGACACGACCACTGGGACTGAGACAG 318  
301 GCCCAGACTCTTACCGGAGGACGACAGTGGGAAATTTTGGACATGGGCGCAAGCCTGATC 360  
319 GCCCAGACTCTTACCGGAGGACGACAGTGGGAAATTTTGGACAAACGGGCCAAGGCTGATC 378  
361 CAGCAATGCGCGGTGAGTGAATAAAGGCTTGGGTTGTAAACCTCTTTCAGTGCAGAAAGA 420  
379 CAGCAATGCGCGGTGAGTGAATAAAGGCTTGGGTTGTAAACCTCTTTCAGTGCAGAAAGA 438  
421 AAAGGTTACGATTAATATCGTGAACCATGACGCGTATCGACAGAAAGACACCGGCTAAC 480  
439 AAAGGTTACGATTAATATCGTGAACCATGACGCGTATCGACAGAAAGACACCGGCTAAC 498  
481 TACGTGCCAGCAGCGCGGTAAATACGTAGAGGTGCAGACGTTAATCGGAATTACTGGGCGT 540  
499 TACGTGCCAGCAGCGCGGTAAATACGTAGAGGTGCAGACGTTAATCGGAATTACTGGGCGT 558  
541 AAAGGTTACGATTAATATCGTGAACCATGACGCGTATCGACAGAAAGACACCGGCTAAC 600  
559 AAAGGTTACGATTAATATCGTGAACCATGACGCGTATCGACAGAAAGACACCGGCTAAC 618  
601 TGGGTTTGAACCTACAAAGCTAGAGTGTGACAGAGGGGGGTGGAATTCCATGTGTAGCAG 660  
619 TGGGTTTGAACCTACAAAGCTAGAGTGTGACAGAGGGGGGTGGAATTCCATGTGTAGCAG 678  
721 TGAATATGCTAGATATGAAAGACATCGATGCGACAGGACGCTCTTGGGTTAACT 720  
679 TGAATATGCTAGATATGAAAGACATCGATGCGACAGGACGCTCTTGGGTTAACT 738  
721 GAGGCTCATGACGAAAGCGTGGGAGCAAAACGAGTTTGAATACCTGTGATGCCAGCC 780  
739 GAGGCTCATGACGAAAGCGTGGGAGCAAAACGAGTTTGAATACCTGTGATGCCAGCC 798  
781 CTAAAGATGTCAACTAGTTTGGGCGCTTATTAGGCTTGTATACGAGCTAACGCGTGA 840

Db	799	CTAAACGATGTCACCTAGTGTGGGGCCCTTACTAGGCTTGGTAAGCTAAGCTAAACGCGTGA	858
Qy	841	AGTTGACCGCGCTGGGAGTACCGGTGCAAGATTAAACTCAAGGAATTGACGGGGAACC	900
Db	859	AGTTGACCGCGCTGGGAGTACCGGTGCAAGATTAAACTCAAGGAATTGACGGGGAACC	918
Qy	901	GCAACAAGCGGTGGATTATGTGGATTAAATTCGATGCAAGCGGAAAACTTACCTTACCCTT	960
Db	919	GCACAAAGCGGTGGATTATGTGGATTAAATTCGATGCAAGCGGAAAACTTACCTTACCCTT	978
Qy	961	GACATGTAGCGCAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTAACACAGGTGCT	1019
Db	979	GACATGTAGCGCAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTAACACAGGTGCT	1038
Qy	1020	GCATGGCTGTGCTGACCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAAGAGCGCAAC	1079
Db	1039	GCATGGCTGTGCTGACCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAAGAGCGCAAC	1098
Qy	1080	CCTTGTCAATTAAATTTGCCATCATTTTGGTGGGCACTTTAATGAGACTGCGGTGACAAAC	1139
Db	1099	CCTTGTCAATTAAATTTGCCATCATTTTGGTGGGCACTTTAATGAGACTGCGGTGACAAAC	1158
Qy	1140	GGAGGAAGGTGGGAGATGACGTAAGTCTCAATGCGCCCTTATGGGAGGGCTTGCACAGTA	1199
Db	1159	GGAGGAAGGTGGGAGATGACGTAAGTCTCAATGCGCCCTTATGGGAGGGCTTGCACAGTA	1218
Qy	1200	ATACAAATGCGCGGTACAGAGGGTTGCGCAACCGCGAGGGGGAGCTAATCTCAGAAAGCGC	1259
Db	1219	ATACAAATGCGCGGTACAGAGGGTTGCGCAACCGCGAGGGGGAGCTAATCTCAGAAAGCGC	1278
Qy	1260	GTCGTAGTCCGGATGGAAGTCTGCAACTGCACTCCGTGAAGTCCGAATGCGTAAATGCG	1319
Db	1279	GTCGTAGTCCGGATGGAAGTCTGCAACTGCACTCCGTGAAGTCCGAATGCGTAAATGCG	1338
Qy	1320	CGGATCAGCATGTCCGGGTGAATTAAGTTCGGGGGTCTGTACACACGCGCGGTGACACCA	1379
Db	1339	CGGATCAGCATGTCCGGGTGAATTAAGTTCGGGGGTCTGTACACACGCGCGGTGACACCA	1398
Qy	1380	TGGAGTGGGTTTCCACGAAGCAGGTAGTCTTAACCGTAAGAGAGGCGCTTGCCACGGTG	1439
Db	1399	TGGAGTGGGTTTCCACGAAGCAGGTAGTCTCTAACCCTCGGGAGGACGCTTACCAACGGTG	1458
Qy	1440	AGATTCAATGACTGGGG 1455	
Db	1459	TGGTCAATGACTGGGG 1474	
RESULT 7			
ADM32721			
ID	ADM32721	standard; DNA; 1467 BP.	
XX	AC	ADM32721;	
XX	DT	17-JUN-2004 (first entry)	
XX	DE	AOB P4clone42 16S rDNA.	
XX	KM	16S rDNA; ammonia oxidizing bacteria; AOB; ammonia; nitrite;	
XX	OS	aqueous environment; freshwater; seawater; aquarium; ss.	
XX	PN	Nitrosomonas aestuarii.	
XX	XX	MO2004026772-A2.	
XX	PD	01-APR-2004.	
XX	PF	10-SEP-2003; 2003WO-US028210.	
XX	PR	19-SEP-2002; 2002US-0386219P.	
XX	PR	19-SEP-2002; 2002US-0386219P.	
XX	PA	(AQUA-) AQUARIA INC.	

XX	Hovanec TA;
F1	WPI; 2004-304936/28.
DR	
XX	New composition comprising an isolated bacterial strain that oxidizes
XX	ammonia to nitrite, useful for alleviating or preventing the accumulation
PT	of ammonia in aqueous environment.
PS	
XX	Claim 1; Page 13, 98pp; English.
CC	This sequence represents a 16S rDNA sequence derived from an ammonia
CC	oxidizing bacteria (AOB). This sequence may be used in a composition
CC	which comprises an isolated bacterial strain that oxidizes ammonia to
CC	nitrite. The composition may be used for alleviating or preventing the
CC	accumulation of ammonia in a medium. The ammonia is reduced by at least
CC	30% when compared with a level of ammonia that would exist in the absence
CC	of the bacterial strain. The composition is useful for alleviating or
CC	preventing the accumulation of ammonia in aqueous environment, e.g. a
SQ	freshwater or seawater aquarium.
	Sequence 1467 BP; 384 A; 320 C; 451 G; 312 T; 0 U; 0 Other;
Query Match	93.2%; Score 1357.2; DB 12; Length 1467;
Best Local Similarity	97.2%; Pred. No. 0; Mismatches 38; Indels 3; Gaps 3
Matches 1413; Conservative	0; Mismatches
OY	1 ATTGAACTGCTGGCGCATGCTTTACATGCAGTGCAGCAGCACGATGCTTGCAAT 60
Db	16 ATTGAAGCCTGGCGCATGCTTTACATGCAGTGCAGCAGCACGCGGTCTTGCAAC 75
OY	61 CTGTGCGCAGTGGCGGACGCGGTGATGATGCATTCGAAACGTATCCAGAAGCGGGGTA 120
Db	76 CTGGTGGCGAATGGCGCGACGCGGTGATGATGCATTCGAAACGTATCCAGAAGCGGGGTA 135
OY	121 ACCGATCGAAGATGTGCTAATACCAGCATATCTCTAAGAGGAAAGCAGGGGATCGAAA 180
Db	136 ACGATCGAAGATGTGCTAATACCAGCATATCTCTAAGAGGAAAGCAGGGGATCGAAA 195
OY	181 GACCTTGGCTTTTGGAGCGGCCGATGCTGATTAGTAGTTGGTGGGGTAAAGGCTTAC 240
Db	196 GACCTTGGCTTTTGGAGCGGCCGATGCTGATTAGTAGTTGGTGGGGTAAAGGCTTAC 255
OY	241 CAAGCGCAGCATCAGTAGTTGGTGTGAGAGGACGACGACCACTGSGACTGAGACACG 300
Db	256 CAAGCGCAGCATCAGTAGTTGGTGTGAGAGGACGACGACCACTGSGACTGAGACACG 315
OY	301 GCCCGACATCTCTAAGGAGGACGACGATGGGGAATTTTGGACAATGGGGCGAACCTGATC 360
Db	316 GCCCGACATCTCTAAGGAGGACGACGATGGGGAATTTTGGACAATGGGGCGAACCTGATC 375
OY	361 CAGCAATGCGCGTGAAGTGAAGAAAGGCTTGGGGTGTAAAGCTCTTCAGTCGAGAAGA 420
Db	376 CAGCAATGCGCGTGAAGTGAAGAAAGG -CTTGGGGTGTAAAGCTCTTCAGTCGAGAAGA 434
OY	421 AAAGGTTACGGTAAATAATTCGTGACCCATGACGATATGACAGAAAGAACCGGCTAAC 480
Db	435 AAAGGTTGATCTAATATATCACACTTATGATGGTACCGACAGAAAGAACCGGCTAAC 494
OY	481 TAGCTGCGACAGCCGCGGTATATACGTAGGGTGCAGAGCTTATATGGAAATTACTGGGCT 540
Db	495 TAGCTGCGACAGCCGCGGTATATACGTAGGGTGCAGAGCTTATATGGAAATTACTGGGCT 554
OY	541 AAAGGTTGCGAGGCGGCTTGTAGTACAGATGGGAATCCCGGGCTTAACTCTGGGAAT 600
Db	555 AAAGGTTGCGAGGCGGCTTGTAGTACAGATGGGAATCCCGGGCTTAACTCTGGGAAT 614
OY	601 TGCGTTTGAACCTCAAAGCTTAGAGTGGCGAGAGGAGTGGATTCATGTGTAGCAG 660
Db	615 TGCGTTTGAACCTCAAAGCTTAGAGTGGCGAGAGGAGGAGTGGATTCATGTGTAGCAG 674
OY	661 TGAATGCGTAGATATGAAAGACATGCATGGCGAAGGACCTCTCTGGGTTAACT 720

Db 675 TGAATGCGTAGAGATATGGAAGAACTGATGCGAGCGACCCCTGGTTAACT 734  
QY 721 GAGCTCATGACAGAAAGCGTGGGAGCAAAAGATTAAGTACCTGTAGTCAAGCC 780  
Db 735 GAGCTCATGACAGAAAGCGTGGGAGCAAAAGATTAAGTACCTGTAGTCAAGCC 794  
QY 781 CTAAACGATGCACTAATGTTGGGCTTATTAGCTGTAAAGAACTCAAGCTGA 840  
Db 795 CTAAACGATGCACTAATGTTGGGCTTATTAGCTGTAAAGAACTCAAGCTGA 854  
QY 841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTTAACTCAAGGAGGAGCC 900  
Db 855 AGTTGACCGCTGGGAGTACGCTCGCAAGATTTAACTCAAGGAGGAGCC 914  
QY 901 GCAAGGCGGTGATTTATGATTAATTAATGATGCAAGGAAAACTTAACCTT 960  
Db 915 GCAAGGCGGTGATTTATGATTAATTAATGATGCAAGGAAAACTTAACCTT 974  
QY 961 GACATGAGCGAATTTCTAGAGATAGATTAGTGC-TTCGGGAAAGCTTAACAGAGTGT 1019  
Db 975 GACATGAGCGAATTTCTAGAGATAGATTAGTGC-TTCGGGAAAGCTTAACAGAGTGT 1034  
QY 1020 GCATGCTGTGCTGACCTGCTGTGTGAGATGTTGGGTTAAGTCCGCAACGAGCGGAC 1079  
Db 1035 GCATGCTGTGCTGACCTGCTGTGTGAGATGTTGGGTTAAGTCCGCAACGAGCGGAC 1094  
QY 1080 CCTTGATTAATTAATGATGATTTGTTGGGCACTTAATGAGACTCGGCTGACAAAC 1139  
Db 1095 CCTTGATTAATTAATGATGATTTGTTGGGCACTTAATGAGACTCGGCTGACAAAC 1154  
QY 1140 GAGAGAGAGTGGGATGACGTCAAGTCTCAAGGCGCTTATGGGTAGGGCTTCAACGTA 1199  
Db 1155 GAGAGAGAGTGGGATGACGTCAAGTCTCAAGGCGCTTATGGGTAGGGCTTCAACGTA 1214  
QY 1200 ATACAAATGAGCGGATGACAGAGGTTGCCAACCCGGAAGGGAGCTTAATCTCAAGAAAGGC 1259  
Db 1215 ATACAAATGAGCGGATGACAGAGGTTGCCAACCCGGAAGGGAGCTTAATCTCAAGAAAGGC 1274  
QY 1260 GTGCTAGTCCGAGTCCGAGTGCAGACTCGAGTCAAGTCCGAAATCGCTAGTAATCG 1319  
Db 1275 GTGCTAGTCCGAGTCCGAGTGCAGACTCGAGTCAAGTCCGAAATCGCTAGTAATCG 1334  
QY 1320 CGATGACGATGTCGCGGTGAATAGTTCGCGGCTTTTGAACACCGCGCTGACACCA 1379  
Db 1335 CGATGACGATGTCGCGGTGAATAGTTCGCGGCTTTTGAACACCGCGCTGACACCA 1394  
QY 1380 TGGAGTGGGTTTCAACGAGAGCAAGTGTCTTAACCTTAAGAGAGGCGCTTCCACGGTG 1439  
Db 1395 TGGAGTGGGTTTCAACGAGAGCAAGTGTCTTAACCTTAAGAGAGGCGCTTCCACGGTG 1453  
QY 1440 AGATTCAATGACTGG 1453  
Db 1454 AGATTCAATGACTGG 1467

RESULT 8  
ABAO2418  
ID ABAO2418 standard; DNA; 1458 BP.

XX ABAO2418;  
AC 29-AUG-2003 (revised)  
DT 04-MAR-2002 (first entry)

XX Type B ammonia-oxidizing bacterium 16S rRNA gene sequence. R3clone5.  
XX Type B; ammonia-oxidizing bacterium; AOB; nitrite; 16S rRNA gene;  
KW ribosomal RNA; aquarium; aquaculture; waste water treatment;  
XX bioremediation; ds.  
XX Nitrosomonadales.  
OS  
PN

WO200190312-A1.

XX 29-NOV-2001.  
PD 17-MAY-2001; 2001WO-US016265.  
XX 19-MAY-2000; 2000US-00573684.  
XX (AQUA-) AQUARIA INC.  
XX Hovanec TA, Burrell PC;  
XX WPI; 2002-075367/10.  
PT New bacteria capable of oxidizing ammonia to nitrite, for preventing or  
PT alleviating the accumulation of ammonia in fresh water aquaria, seawater  
PS aquaria and waste water.  
PS Claim 2; Page 7; 62pp; English.  
XX The invention relates to 4 novel types of ammonia-oxidizing bacteria  
XX (AOB) found in freshwater aquaria. The bacteria are able to oxidise  
XX ammonia to nitrite and are members of the ammonia-oxidizing bacteria  
XX family of the beta subdivision of Proteobacteria. The 4 types of bacteria  
XX can be distinguished on the basis of their 16S rRNA (ribosomal RNA) gene  
XX sequences (ABAO2418-ABAO2419), and are classified as AOB type A (e.g.,  
XX R3clone140), type A1 (e.g., R3clone187), type B (e.g., R3clone5) and type  
XX C (e.g., R3clone47). The invention also encompasses isolated 16S rRNA  
XX gene sequences of the ammonia-oxidizing bacteria of the invention,  
XX oligonucleotide probes and primers for the detection of these bacteria,  
XX and compositions comprising the bacteria. The bacteria of the invention  
XX are useful in biological filters for reducing ammonia accumulation in  
XX both freshwater and seawater aquaria. They may also be used in waste  
XX water treatment and in bioremediation processes to reduce the level of  
XX pollution caused by ammonia. The present sequence represents R3clone5, a  
XX 16S rRNA gene sequence from the type B ammonia-oxidizing bacterium of the  
XX invention. (Updated on 29-AUG-2003 to standardise OS field)  
SQ Sequence 1458 BP; 368 A; 342 C; 466 G; 282 T; 0 U; 0 Other;  
Query Match 88.0%; Score 1282.8; DB 6; Length 1458;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 1363; Conservative 0; Mismatches 92; Indels 5; Gaps 2;  
QY 1 ATGAAGCTGCGCGGATGCTTTTCAATGCAAGTCAAGCGGAGCGAGTGGTGGAT 60  
Db 1 ATGAAGCTGCGCGGATGCTTTTCAATGCAAGTCAAGCGGAGCGAGTGGTGGAT 58  
QY 61 CTGTTGCGAGTGGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 59 CTGTTGCGAGTGGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 118  
QY 121 ACGCATGAAAGATGCTATATACCGCATATCTTAAGAGAGAAAGCGAGGATTCGAA 180  
Db 119 ACGCATGAAAGATGCTATATACCGCATATCTTAAGAGAGAAAGCGAGGATTCGAA 178  
QY 181 GACCTTGGCTTTTGAAGCGGCGCATGCTGATTAAGTGAAGTGGTGGTGAAGGCTTAC 240  
Db 179 GACCTTGGCTTTTGAAGCGGCGCATGCTGATTAAGTGAAGTGGTGGTGAAGGCTTAC 238  
QY 241 CAAAGCGAGATGCTAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
Db 239 CAAAGCGAGATGCTAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 298  
QY 301 GCCCAGACTCTTAAGGAG 360  
Db 299 GCCCAGACTCTTAAGGAG 358  
QY 361 CAGCATGCGCGGATGAG 420  
Db 359 CAGCATGCGCGGATGAG 418  
QY 421 AAAGTTTACGTTAATATGCTGACCATGACGAGTATGACAGAGAGAGAGAGAGAGAG 480

```
Db 419 AACGTCACGGCTAATACCCGTCACCTACGAGTACCGGAAAGAAAGCAACCGGCTAAC 478
Qy 481 TAGCTGCACAGCCGGGCTAATCTAGAGGTGCAACCGTTAATCGGAATTAATCTGGGCGT 540
Db 479 TAGCTGCACAGCCGGGCTAATCTAGAGGTGCAACCGTTAATCGGAATTAATCTGGGCGT 538
Qy 541 AAAGGTCGCGCAGCGCGCTTGTAAAGTCAAGTGAATATCCCGGGCTTAACTGGGAAT 600
Db 539 AAAGCGTCGCGCAGCGCGCTTGTAAAGTCAAGTGAATATCCCGGGCTTAACTGGGAAT 598
Qy 601 TCGCTTTGAAACTACAAAGCTAGAGTGTGCGAGAGGAGGTGGAATTCATGTGTAGCAG 660
Db 599 TCGCTTTGAAACTACAAAGCTAGAGTGTGCGAGAGGAGGTGGAATTCATGTGTAGCAG 658
Qy 661 TGAATTCGCTAGAGATATGGAAGAACTCGATGCGGAAAGGACGCTCTGGGTTAACTACT 720
Db 659 TGAATTCGCTAGAGATATGGAAGAACTCGATGCGGAAAGGACGCTCTGGGTTAACTACT 718
Qy 721 GACGCTCATGCAAGAAAGCGTGGGGAGCAAAACAGAGATTAGATACCTGTAGTCCAGCC 780
Db 719 GACGCTCATGCAAGAAAGCGTGGGGAGCAAAACAGAGATTAGATACCTGTAGTCCAGCC 778
Qy 781 CTAAACGATGTCAACTAGTTGTGGGCTTATTAAGCTTGTGTAACGAAGCTAAACGCGTGA 840
Db 779 CTAAACGATGTCAACTAGTTGTGGGCTTATTAAGCTTGTGTAACGAAGCTAAACGCGTGA 838
Qy 841 AGTTGACCCGCTGCGGAGTACGCTCGCAAGATTAAAACTCAAGAAATTAACGGGGACCC 900
Db 839 AGTTGACCCGCTGCGGAGTACGCTCGCAAGATTAAAACTCAAGAAATTAACGGGGACCC 898
Qy 901 GCAACAAGCGGTGATTTATGTGATTTAATGATGCAACGCGAAACCTTAACCTTACCTT 960
Db 899 GCAACAAGCGGTGATTTATGTGATTTAATGATGCAACGCGAAACCTTAACCTTACCTT 958
Qy 961 GACATGTAGGAAATTTCTAGAGATAGATTAGTGC--TTGCGGAAACGCTTACACAGGTG 1017
Db 959 GACATGTAGGAAACCGCCGCGAGAGGTGGGTGGCCGAAAGGAGCGTAAACACAGGTG 1018
Qy 1018 CTGCAATGCGTGTGTCACGCTCGTGTGCGTGAAGTGTGGTTAAAGTCCCGCAACGAGCGCA 1077
Db 1019 CTGCAATGCGTGTGTCACGCTCGTGTGCGTGAAGTGTGGTTAAAGTCCCGCAACGAGCGCA 1078
Qy 1078 ACCCTGTGATTAATGATGCAATCTTGTGGGCACTTAAATGAGATGCGCGGTGACAA 1137
Db 1079 ACCCTGTGATTAATGATGCAATCTTGTGGGCACTTAAATGAGATGCGCGGTGACAA 1138
Qy 1138 CCGGAGGAGAGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGAGTGGGCTTCAACG 1197
Db 1139 CCGGAGGAGAGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGAGTGGGCTTCAACG 1198
Qy 1198 TAAATCAATGCGCGCTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAAGAAAGC 1257
Db 1199 TAAATCAATGCGCGCTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAAGAAAGC 1258
Qy 1258 GCGTCGATGCGCGGATGCGAGTCTGCAACTGCACTCCGCTGAAGTCGGAATCGCTAGTAAT 1317
Db 1259 GCGTCGATGCGCGGATGCGAGTCTGCAACTGCACTCCGCTGAAGTCGGAATCGCTAGTAAT 1318
Qy 1318 CCGGATCAAGCATGTGCGCGGTGAATAGTCCCGGGCTTGTACACACCGCCCGTCAAC 1377
Db 1319 CCGGATCAAGCATGTGCGCGGTGAATAGTCCCGGGCTTGTACACACCGCCCGTCAAC 1378
Qy 1378 CATGGAGTGGGTTTACCAAGAACAGATAGTCTAAACCGTAAAGAGGGCGCTTGGCAACG 1437
Db 1379 CATGGAGTGGGTTTACCAAGAACAGATAGTCTAAACCGTAAAGAGGGCGCTTGGCAACG 1438
Qy 1438 TGAATTCATGACTGGGGTG 1457
Db 1439 TGAATTCATGACTGGGGTG 1458
```

RESULT 9  
ADM32706

```
ID ADM32706 standard; DNA; 1458 BP.
XX
AC ADM32706;
XX
DT 17-JUN-2004 (first entry)
XX
DE AOB Type B R3clones 16S rDNA.
XX
KW 16S rDNA; ammonia oxidizing bacteria; AOB; ammonia; nitrite;
XX aqueous environment; freshwater; seawater; aquarium; ss.
XX
OS Nitrosomonas sp.
XX
PN W02004026772-A2.
XX
PD 01-APR-2004.
XX
PF 10-SEP-2003; 2003WO-US028210.
XX
PR 19-SEP-2002; 2002US-0386217P.
PR 19-SEP-2002; 2002US-0386218P.
PR 19-SEP-2002; 2002US-0386219P.
XX
PA (AQUA-) AQUARIA INC.
XX
PI Hovanec TA;
XX
DR WPI; 2004-304936/28.
XX
PT New composition comprising an isolated bacterial strain that oxidizes
PT ammonia to nitrite, useful for alleviating or preventing the accumulation
XX of ammonia in aqueous environment.
XX
PS Disclosure; Page 11; 98pp; English.
XX
CC This sequence represents a 16S rDNA sequence derived from an ammonia
CC oxidizing bacteria (AOB). This sequence may be used in a composition
CC which comprises an isolated bacterial strain that oxidizes ammonia to
CC nitrite. The composition may be used for alleviating or preventing the
CC accumulation of ammonia in a medium. The ammonia is reduced by at least
CC 30% when compared with a level of ammonia that would exist in the absence
CC of the bacterial strain. The composition is useful for alleviating or
CC preventing the accumulation of ammonia in aqueous environment, e.g. a
CC freshwater or seawater aquarium.
XX
SQ Sequence 1458 BP; 368 A; 342 C; 466 G; 282 T; 0 U; 0 Other;
XX
Query Match 88.0%; Score 1282.8; DB 12; Length 1458;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1363; Conservative 0; Mismatches 92; Indels 5; Gaps 2;
Qy 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAATGCAAGTCGACGCGAGTCTTGCAAT 60
Db 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAATGCAAGTCGACGCGAGTCTTGCAAT 58
Qy 61 CTGTGCGGAGTGGCGGACGGGTGAGTATGATGCGAAGCGTAATCCAGAAAGGGGGGTA 120
Db 59 CTGTGCGGAGTGGCGGACGGGTGAGTATGATGCGAAGCGTAATCCAGAAAGGGGGGTA 118
Qy 121 ACGCATCGAAGATGTGCTAATACCGCATATCTCTTAAGAGAGAAAGAGGGGATCGAAA 180
Db 119 ACGCATCGAAGATGTGCTAATACCGCATATCTCTTAAGAGAGAAAGAGGGGATCGCAA 178
Qy 181 GACCTTGCGCTTTTGAAGCGCGCGATGTCTGATTAGTATGCTAGTTGGGGTAAAGGCTTAC 240
Db 179 GACCTTGCGCTTTTGAAGCGCGCGATGTCTGATTAGTATGCTAGTTGGGGTAAAGGCTTAC 238
Qy 241 CAAGCGACGATCATAGTGTGTCGAGAGACGACGACGACCACTGGGACTGAGACACG 300
Db 239 CAAGCGACGATCATAGTGTGTCGAGAGACGACGACGACCACTGGGACTGAGACACG 298
Qy 301 GCCGAGACTCTTACGCGGAGGAGCAGAGTGGGGAAATTTTGAACAAATGGGCGCAAGCTGATC 360
Db 301 GCCGAGACTCTTACGCGGAGGAGCAGAGTGGGGAAATTTTGAACAAATGGGCGCAAGCTGATC 360
```

Db 299 GCCCAGACTCTTACGGAGGAGCAGTGGGAAATTTTGACAAATGGGGGAAACCTGATC 358  
Qy 361 CAGCAATGCCCGGTGATGAAAGAGCCCTTGGGTTGTAAGCTCTTTCAGTCGAGAGA 420  
Db 359 CAGCATTGCCCGGTGATGAAAGAGCCCTTGGGTTGTAAGCTCTTTCAGTCGAGAGA 418  
Qy 421 AAAGTTACGGTAAATATATCTGACCCATGACGGTATGACAGAAAGAACCGGCTAAC 480  
Db 419 AAGGTCACGGCTAATACCGGTGACTGACGGTACCGAAAGAAAGCACCGGCTAAC 478  
Qy 481 TACGTGCCAGACCGCGGTAAATACGTAGGGTGCAGAGCTTAATCGAAATTAATCGGGGT 540  
Db 479 TACGTGCCAGACCGCGGTAAATACGTAGGGTGCAGAGCTTAATCGAAATTAATCGGGGT 538  
Qy 541 AAAAGGTGCGGAGCGGCTTGTAAAGTCAATGTGAATCCCGGGCTTAACTCGGCAAT 600  
Db 539 AAAGCTGCGGAGCGGCTTGTAAAGTCAATGTGAATCCCGGGCTTAACTCGGCAAT 598  
Qy 601 TGCCTTTGAAACTACAAAGCTAGAGTGGCAGAGAGGTGGAATTCATGTGTAGAG 660  
Db 599 TGCCTTTGAAACTACAAAGCTAGAGTGGCAGAGAGGTGGAATTCATGTGTAGAG 658  
Qy 661 TGAATGCTTGAATATGAAAGCAATCGATGGCGAAGGCAAGCTCTCGGGTTAAACCT 720  
Db 659 TGAATGCTTGAATATGAAAGCAATCGATGGCGAAGGCAAGCTCTCGGGTTAAACCT 718  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAGAGTATAGTACCTCGTGTAGTCAAGCC 780  
Db 719 GACGCTCATGACGAAAGCGTGGGAGCAAGAGTATAGTACCTCGTGTAGTCAAGCC 778  
Qy 781 CTAACAGATGTCAACTAGTTGTGGCTTATTAAGCTTGTGTAAAGCAAGCTTAACGCTGA 840  
Db 779 CTAACAGATGTCAACTAGTTGTGGCTTATTAAGCTTGTGTAAAGCAAGCTTAACGCTGA 838  
Qy 841 AGTTGACCGCTGGGAGTACGGTGCAGAAATTAAGCTTAAGCAAGATTTGACGGGGACCC 900  
Db 839 AGTTGACCGCTGGGAGTACGGTGCAGAAATTAAGCTTAAGCAAGATTTGACGGGGACCC 898  
Qy 901 GCACAAAGGGTGGATTAATGATTAATTCGATGACGCGGCAAAACCTTACCTACCTT 960  
Db 899 GCACAAAGGGTGGATTAATGATTAATTCGATGACGCGGCAAAACCTTACCTACCTT 958  
Qy 961 GACATGTAGCAATTTTCTAGAGATAGATTAGTGC---TTCGGAAACGCTTAACAGAGGTG 1017  
Db 959 GACATGTAGCAATTTTCTAGAGATAGATTAGTGC---TTCGGAAACGCTTAACAGAGGTG 1018  
Qy 1018 CTGCATGCGTGTCTGACGCTCGTGTCTGAGATGTTGGGTTAAAGTCCCGCAACGAGCCCA 1077  
Db 1019 CTGCATGCGTGTCTGACGCTCGTGTCTGAGATGTTGGGTTAAAGTCCCGCAACGAGCCCA 1078  
Qy 1078 ACCCTGTCAATTAATTTGCATCATTTGGTGGGCACTTAAATAGACCTGCGGTACAAA 1137  
Db 1079 ACCCTGTCAATTAATTTGCATCATTTGGGCACTTAAATAGACCTGCGGTACAAA 1138  
Qy 1138 CCGAGAGAAAGTGGGAGTACGTCAGTCCATGAGCCCTTAAATGAGGCTTACACAG 1197  
Db 1139 CCGAGAGAAAGTGGGAGTACGTCAGTCCATGAGCCCTTAAATGAGGCTTACACAG 1198  
Qy 1198 TAAATCAATGCGCGTACAGAGGTTGGCAACCCCGAGAGGGAGTAAATCTCAGAAAGC 1257  
Db 1199 TAAATCAATGCGCGTACAGAGGTTGGCAACCCCGAGAGGGAGTAAATCTCAGAAAGC 1258  
Qy 1258 GCGTGTAGTCCGGATCCGAGCTGCAACTCGAATCCGGAAGTCCGGAATCCGCTAGTAAT 1317  
Db 1259 GCGTGTAGTCCGGATCCGAGCTGCAACTCGAATCCGGAAGTCCGGAATCCGCTAGTAAT 1318  
Qy 1318 CCGGATCAAGATGTGCGCGGTGAATGCTTCCCGGGTCTTGTATACACACCGCCGCTACAC 1377  
Db 1319 CCGGATCAAGATGTGCGCGGTGAATGCTTCCCGGGTCTTGTATACACACCGCCGCTACAC 1378  
Qy 1378 CATGGAGTGGGTTTACAGAAAGCTGTAACGTTAAAGAGAGGGGCTTTGCCACGG 1437  
Db 1379 CATGGAGTGGGTTTACAGAAAGCTGTAACGTTAAAGAGAGGGGCTTTGCCACGG 1438

Qy 1438 TGAGATTCACTACTGGGGTG 1457  
Db 1439 TGAGATTCACTACTGGGGTG 1458

RESULT 10  
ID ABA02419  
AB 29-AUG-2003 standard; DNA; 1460 BP.  
XX  
AC ABA02419;  
XX  
DT 29-AUG-2003 (revised)  
DT 04-MAR-2002 (first entry)  
XX  
DE Type C ammonia-oxidising bacterium 16S rRNA gene sequence, R3clone47.  
XX  
KW Type C; ammonia-oxidising bacterium; AOB; nitrite; 16S rRNA gene;  
KW ribosomal RNA; aquarium; aquaculture; waste water treatment;  
KW bioremediation; ds.  
XX  
OS Nitrosomonadaceae.  
XX  
PN NC0200190312-AL.  
XX  
PD 29-NOV-2001.  
XX  
PF 17-MAY-2001; 2001WO-US016265.  
XX  
PR 19-MAY-2000; 2000US-00573684.  
XX  
PA (AQUA-) AQUARIA INC.  
XX  
PI Hovaneq TA, Burrell PC;  
XX  
DR WPI; 2002-075367/10.  
XX  
PT New bacteria capable of oxidizing ammonia to nitrite, for preventing or  
PT alleviating the accumulation of ammonia in fresh water aquaria, seawater  
PT aquaria and waste water.  
XX  
PS Claim 2; Page 8; 62pp; English.  
XX  
CC The invention relates to 4 novel types of ammonia-oxidising bacteria  
CC (AOB) found in freshwater aquaria. The bacteria are able to oxidise  
CC ammonia to nitrite and are members of the ammonia-oxidising bacteria  
CC family of the beta subdivision of Proteobacteria. The 4 types of bacteria  
CC can be distinguished on the basis of their 16S rRNA (ribosomal RNA) gene  
CC sequences (ABA02416-ABA02419), and are classified as AOB type A (e.g.,  
CC R3clone140), type A1 (e.g., R3clone187), type B (e.g., R3clone5) and type  
CC C (e.g., R3clone47). The invention also encompasses isolated 16S rRNA  
CC gene sequences of the ammonia-oxidising bacteria of the invention,  
CC oligonucleotide probes and primers for the detection of these bacteria,  
CC and compositions comprising the bacteria. The bacteria of the invention  
CC are useful in biological filters for reducing ammonia accumulation in  
CC both freshwater and seawater aquaria. They may also be used in waste  
CC water treatment and in bioremediation processes to reduce the level of  
CC pollution caused by ammonia. The present sequence represents R3clone47, a  
CC 16S rRNA gene sequence from the type C ammonia-oxidising bacterium of the  
CC invention. (Updated on 29-AUG-2003 to standardise OS field)  
XX  
SQ Sequence 1460 BP; 388 A; 316 C; 454 G; 302 T; 0 U; 0 Other;

Query Match 84.1%; Score 1225.4; DB 6; Length 1460;  
Best Local Similarity 92.1%; Pred. No. 0;  
Matches 1348; Conservative 0; Mismatches 106; Indels 9; Gaps 5;

Qy 1 ATTGAACCTGGCGGATGCTTTTACATGCAATGCAAGCGGACGACGATGCTTGCAT 60  
Db 1 ATTGAACCTGGCGGATGCTTTTACATGCAATGCAAGCGGACGACGATGCTTGCAT 58  
Qy 61 CTGCTGGGAGTGGGCGGAGCGGTGATGATGCAAGCGTATCCAGAAAGAGGGGGGTA 120

Db 59 CTGCCGGCGATGGCGCAACGGGTGATGATATACATCGAACGTGCTTAAAGTGGGGATA 118  
Qy 121 ACCGATGCAAAAGATGCTCTAATACCGGATATCTCTAAGGAAAGCAAGGATCGAAA 180  
Db 119 ACCGATGCAAAAGATGCTCTAATACCGGATAT-CTCTAAGGAAAGCAAGGATCGAAA 177  
Qy 181 GACCTTGCGCTTTTGGAGCGGCGGATGCTGATTAAGTGGTGGGTAAAGCCTTAC 240  
Db 178 GACCTTGCGCTTAAAGAGCGGCGGATGCTGATTAAGTGGTGGGTAAAGCCTTAC 237  
Qy 241 CAAGCGCAGCATGATGTTGCTGAGAGACGACGACGACGACGACGACGACGACGACG 300  
Db 238 CAAGCGCAGCATGATGTTGCTGAGAGACGACGACGACGACGACGACGACGACGACG 297  
Qy 301 GCCCAAGCTCTTACGAGGAGCAGCAGTGGGAAATTTTGGACAATGGGCGCAACCTGATC 360  
Db 298 GCCCAAGCTCTTACGAGGAGCAGCAGTGGGAAATTTTGGACAATGGGCGCAACCTGATC 357  
Qy 361 CAGCAATGCCGCGTGAAGAGAGGCTTCCGGTTTAAAGCTCTTTCAGTGAAGAGA 420  
Db 358 CAGCAATGCCGCGTGAAGAGAGGCTTCCGGTTTAAAGCTCTTTCAGTGAAGAGA 417  
Qy 421 AAAGGTTACGGTAAATTAATCGTACCCATGACGGTATCGACAGAAAGACCGGCTTAC 480  
Db 418 AAAGTATCATGATGATTAATTAATGATTAATGACGATCTGACAGAAAGACCGGCTTAC 477  
Qy 481 TAGCTGCGACGACCGCGGCTTAATACGTAAGGCTGACGACGCTTAACTGAGGCGCT 540  
Db 478 TAGCTGCGACGACCGCGGCTTAATACGTAAGGCTGACGACGCTTAACTGAGGCGCT 537  
Qy 541 AAAGGTTACGGGCGGCGCTTGAAGTCAATGTAAGTCAATGTAAGTCAATGTAAGTCA 600  
Db 538 AAAGGTTACGGGCGGCGCTTGAAGTCAATGTAAGTCAATGTAAGTCAATGTAAGTCA 597  
Qy 601 TGGGTTTGAATCAAAAGCTAAGTGTGCAAGGAGGTGAATTCATGTGTAGCAG 660  
Db 598 TGGGTTTGAATCAAAAGCTAAGTGTGCAAGGAGGTGAATTCATGTGTAGCAG 657  
Qy 661 TGAATGCGCTAAGATTAAGAAACATCGATGGCGGACGCTTCCGCTTAACTGAT 720  
Db 658 TGAATGCGCTAAGATTAAGAAACATCGATGGCGGACGCTTCCGCTTAACTGAT 717  
Qy 721 GACGCTCATGCAAGAAAGCTGGGAGCAAAACAGATTAATGATACCTGATGATCAGC 780  
Db 718 GACGCTCATGCAAGAAAGCTGGGAGCAAAACAGATTAATGATACCTGATGATCAGC 777  
Qy 781 CTAAACGATGTCAAGTGTGTTGGGCTTATTA--GGCTTGTGAAGAACTTACGCGT 838  
Db 778 CTAAACGATGTCAAGTGTGTTGGGCTTATTAAGGATTTGGTAACTGATGATGCGGT 837  
Qy 839 GAAGTTGACCGGCTGGGAGATACGGTGGCAAGATTAATACTCAAAAGAAATTTGACGGGAC 898  
Db 838 GAAGTTGACCGGCTGGGAGATACGGTGGCAAGATTAATACTCAAAAGAAATTTGACGGGAC 897  
Qy 899 CGCACAAGAGGAGTATGATGATTAATTCGATGCAAGCAAGCAAGCAAGCAAGCAAGCA 958  
Db 898 CGCACAAGAGGAGTATGATGATTAATTCGATGCAAGCAAGCAAGCAAGCAAGCAAGCA 957  
Qy 959 TTGACATGTAGCGAATTTCTAGAGATGATTAAGTGC--TTGCGGAACGCTTACACAG 1015  
Db 958 TTGACATGTAGCGAATTTCTAGAGATGATTAAGTGC--TTGCGGAACGCTTACACAG 1017  
Qy 1016 TGTGCAATGGCTGTGCTGAGCTGTGCTGAGAAATGTTGGGTTAATGCTCCGCAACGAGC 1075  
Db 1018 TGTGCAATGGCTGTGCTGAGCTGTGCTGAGAAATGTTGGGTTAATGCTCCGCAACGAGC 1077  
Qy 1076 CAACCTTGTCTAATTAATGCTCATCAAT--TGTGTTGGGCACTTATGATGATGCTGCGGTGAC 1134  
Db 1078 CAACCTTGTCTAATTAATGCTCATCAATTAATGATGATGATGATGATGATGATGATGAT 1137  
Qy 1135 AAACCGAAGAGAGTGGAGATGACGTCAGTCTCATGAGCCCTTATGAGGAGCTTAC 1194  
Db 1138 AAACCGAAGAGAGTGGAGATGACGTCAGTCTCATGAGCCCTTATGAGGAGCTTAC 1197

Qy 1195 AGCTAATACAAATGGCGGCTGACAGAGGTTTCCCAACCGCGAGAGGAGCTAATCTCAGAA 1254  
Db 1198 AGCTAATACAAATGGCGGCTGACAGAGGTTTCCCAACCGCGAGAGGAGCTAATCTCAGAA 1257  
Qy 1255 AGCGGCTGTAGTCCGGAATCGAAGTCTGCACTGCTCGTGAATCGGAATCGCTAGT 1314  
Db 1258 AGCGGCTGTAGTCCGGAATCGAAGTCTGCACTGCTCGTGAATCGGAATCGCTAGT 1317  
Qy 1315 AATCGGAGATCAGCATGTCGCGGCTGAAATACGTTCCCGGCTCTTGTACACACCGCGCTGA 1374  
Db 1318 AATCGGAGATCAGCATGTCGCGGCTGAAATACGTTCCCGGCTCTTGTACACACCGCGCTGA 1377  
Qy 1375 CACCATGGAGAGTGGGTTTACCAAGAGCAGTATGATTAACCGTAAAGAGGAGCTTGCCA 1434  
Db 1378 CACCATGGAGAGTGGGTTTACCAAGAGCAGTATGATTAACCGTAAAGAGAGCCTTGCCA 1437  
Qy 1435 CGGTGAGATTCATGACTGGGGTG 1457  
Db 1438 CGGTGAGGAGTCACTGAGCTGGGGTG 1460

RESULT 11  
ADM32707  
ID ADM32707 standard; DNA; 1460 BP.  
XX  
AC ADM32707;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
XX  
DE AOB Type C R5clone47 16S rDNA.  
XX  
KM 16S rDNA: ammonia oxidizing bacteria; AOB: ammonia; nitrite;  
XX  
KM aqueous environment; freshwater; seawater; aquarium; 88.  
XX  
OS Nitrosomonas sp.  
XX  
PN MO2004026772-A2.  
XX  
PD 01-APR-2004.  
XX  
XX  
PF 10-SEP-2003; 2003WO-US028210.  
XX  
PR 19-SEP-2002; 2002US-0386217P.  
PR 19-SEP-2002; 2002US-0386218P.  
PR 19-SEP-2002; 2002US-0386219P.  
XX  
PA (AQUA-) AQUARIA INC.  
XX  
PI Hovanec TA;  
XX  
DR MPI; 2004-304936/28.  
XX  
PT New composition comprising an isolated bacterial strain that oxidizes  
PT ammonia to nitrite, useful for alleviating or preventing the accumulation  
PT of ammonia in aqueous environment.  
XX  
PS Disclosure; Page 12; 98bp; English.  
XX  
XX This sequence represents a 16S rDNA sequence derived from an ammonia  
CC oxidizing bacteria (AOB). This sequence may be used in a composition  
CC which comprises an isolated bacterial strain that oxidizes ammonia to  
CC nitrite. The composition may be used for alleviating or preventing the  
CC accumulation of ammonia in a medium. The ammonia is reduced by at least  
CC 30% when compared with a level of ammonia that would exist in the absence  
CC of the bacterial strain. The composition is useful for alleviating or  
CC preventing the accumulation of ammonia in aqueous environment, e.g. a  
CC freshwater or seawater aquarium.  
XX  
SQ Sequence 1460 BP; 388 A; 316 C; 454 G; 302 T; 0 U; 0 Other;  
Query Match 84.1%; Score 1225.4; DB 12; Length 1460;  
Best Local Similarity 92.1%; Pred. No. 0;

Matches 1348; Conservative 0; Mismatches 106; Indels 9; Gaps 5;

```

Qy 1 ATTTGAACGCTGGCGGCAATGCTTTTACATGCAAGTCGAAACGGCGAGCAGCATGCTTGCAT 60
Db 1 ATTTGAACGCTGGCGGCAATGCTTTTACATGCAAGTCGAAACGGCGAG--CGGGGGCTTGGC 58
Qy 61 CTGGTGGCGAGTGGCGGAGCGGGTGAATGTCATCGGAACGTATCCAGAAAGGGGGGTA 120
Db 59 CTGGCGGCGAGTGGCGGAGCGGGTGAATGTCATCGGAACGTATCCAGAAAGGGGGGTA 118
Qy 121 ACCGATGAAAGATGTCTAATACCGCATATCTCTAAGAGGAAACCAAGGGGATCGAAA 180
Db 119 ACCGATGAAAGATGTCTAATACCGCATAT-CTCTAAGAGGAAACCAAGGGGATCGCAA 177
Qy 181 GACCTTGGCTTTTGGAGCCGCGCATGTCTGATTAGCTAGTTGGTGGGTTAAAGCCCTAC 240
Db 178 GACCTTGGCTTTAAAGAGCCGCGCATGTCTGATTAGCTAGTTGGTGGGTTAAAGCCCTAC 237
Qy 241 CAAGGCGACGATCAGTAGTTGGTCTGAGAGGACGACCACTGGGACTGAGACACG 300
Db 238 CAAGGCGACGATCAGTAGTTGGTCTGAGAGGACGACCACTGGGACTGAGACACG 297
Qy 301 GCCGAGCTCTTACCGGAGGAGCAGATGGGGAAATTTGGACAAATGGGCGCAAGCTGATC 360
Db 298 GCCGAGCTCTTACCGGAGGAGCAGATGGGGAAATTTGGACAAATGGGCGCAAGCTGATC 357
Qy 361 CAGCAATGCGCGGAGTGAAGAAAGGCTTGGGTTGAAAGCTCTTTCAGTCAGAGAGA 420
Db 358 CAGCAATGCGCGGAGTGAAGAAAGGCTTGGGTTGAAAGCTCTTTCAGTCAGAGAGA 417
Qy 421 AAAGGTTACGCTAATATCGTGAACCATGACGCGTATCGACAGAGAAAGACCGGCTAAC 480
Db 418 AAAGGTTACGCTAATATCGTGAACCATGATGATGATGACGAGTACGACAGAAAGACCGGCTAAC 477
Qy 481 TACGTGCGACGACCGCGGCTAATACGTAGGGTGCAGAGGCTTAATCGGAAATTAAGTGGGCT 540
Db 478 TACGTGCGACGACCGCGGCTAATACGTAGGGTGCAGAGGCTTAATCGGAAATTAAGTGGGCT 537
Qy 541 AAAGGTTGCGGAGCGGCTTGTAGTGAATGTAATCCCGGGCTTAACTCGGGGAAAT 600
Db 538 AAAGGTTGCGGAGCGGCTTGTAGTGAATGTAATCCCGGGCTTAACTCGGGGAAAT 597
Qy 601 TGCGTTGAAACTCAAAAGCTAGAGTGGCAGAGGAGGTGAATTCATGTGTAGCAG 660
Db 598 TGCGTTGAAACTCAAAAGCTAGAGTGGCAGAGGAGGTGAATTCATGTGTAGCAG 657
Qy 661 TGAATGCGTAGAGATGTGAAGAACATGATGGCGAAGGCGCTCTGGGTTAACACT 720
Db 658 TGAATGCGTAGAGATGTGAAGAACACGATGGCGAAGGCGCTCTGGGTTGACACT 717
Qy 721 GACGCTGACGACGAAACGCTGGGAGGAAACAGGATAGATACCTGTGTAGTCCAGCC 780
Db 718 GACGCTGACGACGAAACGCTGGGAGGAAACAGGATAGATACCTGTGTAGTCCAGCC 777
Qy 781 CTAAACGATGCACTAGTTGGGCTTATTA--GGCTTGTATACGAAAGCTAACCGCT 838
Db 778 CTAAACGATGCACTAGTTGGGCTTATTAAGTATTAAGATTTGTAAACGTAACGTAACCGCT 837
Qy 839 GAAAGTTGACCGCTGGGAGATACGCTCGCAAGATTTAAACTCAAGAAATTAAGCGGGAC 898
Db 838 GAAAGTTGACCGCTGGGAGATACGCTCGCAAGATTTAAACTCAAGAAATTAAGCGGGAC 897
Qy 899 CCGCACAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGGAAACCTTAACCTACCC 958
Db 898 CCGCACAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGGAAACCTTAACCTACCC 957
Qy 959 TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTC--TTCCGGAACGCTTAACACAG 1015
Db 958 TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTC--TTCCGGAACGCTTAACACAG 1017
Qy 1016 TGGTCATAGGCTGTGCTCAGCTGCTGTGCTGAGATGTTGGGTTAAAGTCCCGCAACAGCG 1075
Db 1018 TGGTCATAGGCTGTGCTCAGCTGCTGTGCTGAGATGTTGGGTTAAAGTCCCGCAACAGCG 1077

```

```

Qy 1076 CAACCTTGTCACTTAATTGCCATCATT-TGTTGGGCACTTTAATGAGACTGCCGGTGAC 1134
Db 1078 CAACCTTGTCACTTAATTGTCTATCACTTAATGAGCACTTTAGTAGACTGCCGGTGAC 1137
Qy 1135 AAACCGGAGAGAGTGGGAGTAGACGTCAAGTCCCTCATGGCCCTTAATGGTAGGGCTTCAC 1194
Db 1138 AAACCGGAGAGAGTGGGAGTAGACGTCAAGTCCCTCATGGCCCTTAATGGTAGGGCTTCAC 1197
Qy 1195 ACCTAATACAAATGCGCGTACAGAGGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAA 1254
Db 1198 ACCTAATACAAATGCGCGTACAGAGGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAA 1257
Qy 1255 AGCGCGTCTGATTCGCCGATGGAGTGTGCAACTCTGACTCCGTAAGTCGGAATCGCTAGT 1314
Db 1258 AGCAACGTCGTATGTCGGAATCGAGTGTGCAACTCTGACTCCGTAAGTCGGAATCGCTAGT 1317
Qy 1315 AATCGCGGATCAGCATGTGCGGTTGAATACGTTCCGGGCTTGTACACACCGCCGTCA 1374
Db 1318 AATCGCGGATCAGCATGTGCGGTTGAATACGTTCCGGGCTTGTACACACCGCCGTCA 1377
Qy 1375 CACCATGAGAGTGGGTTTACAGAAAGCAGTAGTCTTAACCGTAAGAGGGCGCTTGCCA 1434
Db 1378 CACCATGAGAGTGGGTTTACAGAAAGCAGTAGTCTTAACCGTAAGAGAGCGCTTGCCA 1437
Qy 1435 CGGTAGATTCACTGACTGGGGTG 1457
Db 1438 CGGTGGGGTCACTGACTGGGGTG 1460

```

RESULT 12  
AAC86030  
ID AAC86030 standard; cDNA; 1485 BP.  
XX  
AC AAC86030;  
XX  
XX 11-SEP-2003 (revised)  
DT 29-AUG-2001 (first entry)  
XX  
DE 16S rDNA, ProPel.  
XX  
KM 16S rDNA; polyphosphate accumulating organism; PAO; probe; primer;  
KM detection; phosphorus; waste water; sludge; ss.  
OS *Protoniivibrio pelophilus*.  
XX  
PN MO200146459-A1.  
XX  
XX 28-JUN-2001.  
PD  
XX 28-DEC-2000; 2000WO-AU001611.  
PF  
XX 23-DEC-1999; 99AU-0004867.  
PR  
XX (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.  
PA Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ;  
XX  
XX WPI: 2001-408656/43.  
DR  
XX  
XX Novel oligonucleotide probe or primer useful for detecting polyphosphate  
PT accumulating organism in a sample, comprises a sequence that is unique to  
PT 16S rDNA of polyphosphate accumulating organisms.  
XX  
XX  
PS Claim 4; Fig 3; 54p; English.  
XX  
XX The sequences given in AAC86021-30 represent 16S rDNA sequences from  
CC polyphosphate accumulating organisms (PAOs). Sequences which are unique  
CC to these 16S rDNA sequences are used to create a probe or primer for  
CC detecting the relevant organisms. The primer/probe sequences are useful  
CC for detecting PAO cells in a sample, by treating cells in the sample to  
CC fix cellular contents, contacting fixed cells with the primer/probe which  
CC is labelled with a radiolabel, a reporter group or a hapten, under



conditions which allow the probe to hybridize with 16S rRNA within the fixed cell, removing unhybridized probe from the fixed cells, and detecting the labeled probe-RNA hybrid by fluorescence *in situ* hybridization. The primer/probe sequences are useful for identifying PAOs that are capable of biologically removing phosphorus from waste water. Rapid assessment of the presence of a number of PAOs in a waste water sample, can be done using the primer/probe sequences. They allow quick and convenient assessment of whether a sludge or waste water sample includes PAOs and allows quantitation of PAO cells in samples. (Updated on 11-Sep-2003 to standardise OS field)

**SQ** Sequence 1485 BP; 367 A; 341 C; 482 G; 295 T; 0 U; 0 Other;

Query Match	82.5%	Score 1201.6	DB 4	Length 1485
Best Local Similarity	89.8%	Pred No. 0		
Matches 1313	Conservative	0	Mismatches 144	Indels 5
				Gaps 2

Qy	ATTGAAGCGCTGGGGGCGATGCTTTTAAACAATGCAAGTCGAAACGGCAGACAGAAATGCTTCAT	60
Dp	ATTGAAGCGCTGGGGGCGATGCTTTTAAACAATGCAAGTCGAAACGGCAGACAGAAATGCTTCAT	67
Qy	CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGAACGTAATCAGAAAGGGGGGCTA	120
Dp	CTGATGGCGAGTGGCGCAACGGGTGAGTAATGCAATCGAAACGTAATCCCGGAATGGGGGATA	127
Qy	181 GACCTTGCGCTTTTGGAGCGGCGCGAATGTCGATTAGCTAGTTGGTGGGGTTAAAGCCCTAC	240
Dp	188 GACCTCTCGCTTTGGAGCGGCGCGAATGTCGATTAGCTAGTTGGTGGGGTTAAAGCCCTAC	247
Qy	241 CAAGCGACGATCAGTATGTTGGTCTTGAGAGACGACACGCAACTGGGACTTGAGACAG	300
Dp	248 CAAGCGACGATCCGTAAGCGGGCTTGAGAGAGATGATCCGCCAACTGGGACTTGAGACAG	307
Qy	301 GCCCAGACTCTTACGGGAGGACGACGTGGGGAAATTTGGACAAATGGGCGCAAGCTGATC	366
Dp	308 GCCCAGACTCTTACGGGAGGACGACGTGGGGAAATTTGGACAAATGGGCGCAAGCTGATC	367
Qy	361 CAGCAATGCGCGCTGAGTGAAGAAAGGCTTCGGTGTGTAAGCTCTTTCAGTCAGAGAGA	420
Dp	368 CAGCAATGCGCGCTGAGTGAAGAAAGGCTTCGGTGTGTAAGCTCTTTCGGTGGAGAGA	427
Qy	421 AAAGGTTACGGTAAATATATCTGTGACCCATGACGGTATGACAGAAAGACACCGGCTTAC	480
Dp	428 AAATGGCAGCGCTCTTAATATAGCGTGTGTAAGAGGATACCGAATTAAGAAAGCACCGGCTTAC	487
Qy	481 TAGCTGCACGACCGCGGTAATACGTAGAGGTGCAAGCGTTAATCGAATTACTGGGCGT	540
Dp	488 TAGCTGCACGACCGCGGTAATACGTAGAGGTGCAAGCGTTAATCGAATTACTGGGCGT	547
Qy	541 AAAGGTTGCGCAGCGCGGCTTGTAAATCAGATGTGAAATCCCGGGCTTAACTTGGGAAT	600
Dp	548 AAAGCGTGGCAGCGCGGCTTGTGTAAATCAGAGGTGAAATCCCGGGCTCAACTTGGGAAT	607
Qy	601 TGGCTTTGAAATCTAACAAACTAGATGTGTGGCAGAGGAGGTGTGAATTTCAATGTGTACAG	666
Dp	608 GGGCTTTGAAATCTGACCGCTAGATGTGTGACAGAGGAGGATGTGAATTTCAAGTGTACAG	667
Qy	661 TGAATATGCGTAGATATGTAAGAAACATGATGTGGAGGACGCTCTCGGGTTAAACAT	720
Dp	668 TGAATATGCGTAGATATGTAAGAAACATGATGTGGAGGACGCTCTCGGGTTAAACAT	727
Qy	721 GACGCTCATGACGAAAGCGTGGGAGACAAACGATTTAGATACCTGTGTATGTCACGCC	780
Dp	728 GACGCTCATGACGAAAGCGTGGGAGACAAACGATTTAGATACCTGTGTATGTCACGCC	787
Qy	781 CTAAAGATGTCAACTATGTTGTTGG--GCTTATTAAGCTTGTGTAAGAAAGCTAAACGCGT	838
Dp	788 CTAAAGATGTCAACTATGTTGTTGGAGGGTTAAACCTTTAGTGTGTATGCTAAACGCGT	847

QY	839	GAACCTTGACCGCCCTGGGGAGAGTACGGTGTGCAGAAAGATTAAACTCCAAAGAAATTGACGGGGAC	898
Db	848	GAAGTTGACCGCCCTGGGGAGTACCGCCGCAAGGCTTAATACTCAAGAAATTGACGGGGAC	907
QY	899	CCGCAAGAAGCGGTGTGATTAATGTGGATTAAATTCGATGCMAAGCGAAAACTTACTTACC	958
Db	908	CCGCAAGAAGCGGTGTGATTAATGTGGATTAAATTCGATGCMAAGCGAAAACTTACTTACC	967
QY	959	TTGACATGTAGCCGAATTTTCTAGAGATTAGTTAGTC--TTGGGAAACGCTTACACAGG	1015
Db	968	TTGACATGTAGCCGAATTTTCTAGAGATTAGTTAGTTAGTC--TTGGGAAACGCTTACACAGG	1027
QY	1016	TGCGCATATGGCTGTGTCAGCTCGTGTGCGAGATGTTGGAGTTAAGTCCCGCAACGAGCG	1075
Db	1028	TGCGCATATGGCTGTGTCAGCTCGTGTGCGAGATGTTGGAGTTAAGTCCCGCAACGAGCG	1087
QY	1076	CAACCCCTTGTCATTAATTTGCGCATCATTTGGTTGGGCACTTTAAATGAGACTGCGGTGACA	1135
Db	1088	CAACCCCTTGTCATTAATTTGCGCATCATTTAAGTTGGGCACTTTAAATGAGACTGCGGTGACA	1147
QY	1136	AACCGAGAAAGGTGGGGATGACGTCAAGTCTTCATGCGCCCTTAATGGGTAGGGCTTCACA	1195
Db	1148	AACCGAGAAAGGTGGGGATGACGTCAAGTCTTCATGCGCCCTTAATGGGTAGGGCTTCACA	1207
QY	1196	CGTAATACATATGCGCGCTTACAGAGGGTTGGCAACCGCGAGGGGGAGCTAAATCTCAGAA	1255
Db	1208	CGTAATACATATGCGCGCTTACAGAGGGTTGGCAACCGCGAGGGGGAGCTAAATCTCAGAA	1267
QY	1256	GCAGGTGTAGTCCGGATCCGGAATCTGCAACTCGACTCCGTGAAGTCGAAATCGCTAGTA	1315
Db	1268	GCAGGTGTAGTCCGGATCCGGAATCTGCAACTCGACTCGGTGAAGTCGAAATCGCTAGTA	1327
QY	1316	ATCGCGGATCAGCATGTGTCGCGGTGAATACGTTCCGGGTCCTGTATACACACCGCCGTCAC	1375
Db	1328	ATCGCGGATCAGCATGTGTCGCGGTGAATACGTTCCGGGTCCTGTATACACACCGCCGTCAC	1387
QY	1376	ACCATGGGAGTGGGTTTCCACCAAGAGAGTACTTAAACCTGAAGAAGGGCGCTTGCCAC	1435
Db	1388	ACCATGGGAGCGGGTTTTCGCCAAGAGTAAAGTAACTTAACCGCAAGAGAGGGCGCTTACAC	1447
QY	1436	GGTAGAATTCATGACTGGGGTG 1457	
Db	1448	GGCGGGGTTCTGTACTGGGGTG 1469	
RESULT 13			
ABQ78660			
ID	ABQ78660	standard; DNA; 1495 BP.	
XX	AC	ABQ78660;	
XX	DT	25-NOV-2002 (first entry)	
XX	DE	Nucleotide sequence of a fragment of 16S rDNA gene.	
XX	KW	16S rDNA; nitrogen oxide; nitrogen; ss.	
XX	OS	Denitromonas aromaticus.	
XX	PN	JP2002142755-A.	
XX	PD	21-MAY-2002.	
XX	PF	14-NOV-2000; 2000JP-00346844.	
XX	PR	14-NOV-2000; 2000JP-00346844.	
XX	PA	(DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.	
XX	PA	(BIOI-) BIOINDUSTRY KYOKAI SH.	
XX	PA	(AJIN) AJINOMOTO KK.	
XX	WP	WPI; 2002-611517/66.	



PT Microorganism belonging to Denitromonas for removal of nitrogen oxide by  
 PT reduction of nitrogen oxide to nitrogen in gas or aqueous solution, has  
 PT specific mycological characters.

PS Disclosure; Page 6-7; Bpp; Japanese.

XX The present sequence represents a fragment of 16S rDNA gene of  
 CC Denitromonas aromaticus. This is a novel microorganism belonging to the  
 CC new genus Denitromonas, and which has specific mycological characters.  
 CC The new microorganism efficiently reduces and removes nitrogen oxide. The  
 CC bacterium is used for the removal of nitrogen oxide. It reduces nitrogen  
 CC oxide to nitrogen in a gas or aqueous solution

XX Sequence 1495 BP; 363 A; 350 C; 483 G; 299 T; 0 U; 0 Other;

Query Match 81.9%; Score 1194; DB 6; Length 1495;  
 Best Local Similarity 89.5%; Pred. No. 0;  
 Matches 1307; Conservative 0; Mismatches 150; Indels 3; Gaps 2;

```

QY 1 ATTGAAGCTTGGCGGATGCTTTACATGCAATGCAAGCGGACGAGTCTTGCAT 60
DB 16 ATTGAAGCTTGGCGGATGCTTTACATGCAATGCAAGCGGACGAGTCTTGCAT 75
QY 61 CTGGTGGCGAGTGGCGGACGCGGTGATGATCATCGGACGATTCGAGAGGGGGGTA 120
DB 76 CTGGTGGCGAGTGGCGGACGCGGTGATGATCATCGGACGATTCGAGAGGGGGTA 135
QY 121 ACGCATGGAAGATGTCTTAACCGCATTAACCTTAAGAGGAAAGCGGGATCGAAA 180
DB 136 GCCCGGGGAAAGCGGATTAATACCGCATTAATCTTGAGGAGAAAGCGGGGACCTTCG 195
QY 181 GACCTTGGCTTTTGGAGCGGCGCATGCTGATTAGTGTGGGTAAAGCGCTTAC 240
DB 196 GGCCTCGCGCTACTGAGCGGCGGTATGTCGATTAAGTATGGGTAAAGCGCTTAC 255
QY 241 CAAGCGCAAGATCATGATGTTGTTGAGAGACACACGACCACTGGGATCGAGACG 300
DB 256 CAAGCGCAAGATCATGATGTTGTTGAGAGATGATCAAGCACTGGGATCGAGACG 315
QY 301 GCCGAGCTCTTACGGGAGGAGACAGTGGGAAATTTGGACAAATGGGCGCAAGCTGATC 360
DB 316 GCCGAGCTCTTACGGGAGGAGACAGTGGGAAATTTGGACAAATGGGCGCAAGCTGATC 375
QY 361 CAGCAATGCCGCGTGAAGTGAAGAAAGGCTTCGGGTTGTAAGCTCTTCACTCGAGAGA 420
DB 376 CAGCAATGCCGCGTGAAGTGAAGAAAGGCTTCGGGTTGTAAGCTCTTCACTCGAGAGA 435
QY 421 AAGGTTACGGTAATTAATCGTGAACCCATGACGGTATCGACAGAAAGACCGGCTAAC 480
DB 436 AATCATGCGCCCGAATACGGGCGGGGATGACGGTATCATGCAAGAAAGACCGGCTAAC 495
QY 481 TACGTGCCAGACAGCGCGGTGTAATAGTAGGGTGCAGAGGTTAATGGAAATTAATGGGGCT 540
DB 496 TACGTGCCAGACAGCGCGGTGTAATAGTAGGGTGCAGAGGTTAATGGAAATTAATGGGGCT 555
QY 541 AAGGAGTGCAGAGGCGGCTTGTATAGTCAATGTAATCCCGGGCTTAACTTGGGAT 600
DB 556 AAGGAGTGCAGAGGCGGCTTGTGTATAGACAGAGTGTAAATCCCGGGCTTAACTTGGGAA 615
QY 601 TCGCTTTGAAATCTAAGAGTGAAGTGTGCGAGAGGAGTGAATTCATGTGTAGAG 660
DB 616 TCGCTTTGAACTGCAAGGCTTGAAGTGAAGTGAAGGAGGAGTGAATTCATGTGTAGAG 675
QY 661 TGAATTTGGTAGATATATGAGAAATCATGATGGGAGAGGAGGCTCTCGGGTGAACA 720
DB 676 TGAATTTGGTAGATATATGAGAAATCATGATGGGAGAGGAGGCTCTCGGGCTGTACT 735
QY 721 GAGGCTCATGCAAGAAAGGCTGGGAGCAAAAGAGATTAGATACCTGGTAGTCCAGCC 780
DB 736 GAGGCTCATGCAAGAAAGGCTGGGAGCAAAAGAGATTAGATACCTGGTAGTCCAGCC 795
QY 781 CTAAACGATGTCACTAGTTGTT--GGGCTTATTAGGCTTGTGAACGAAAGCTTAACGGCT 838

```

```

DB 796 CTAAACGATGTCACTAGTTGTTCCGAGAGAGAAACTTTCGGGTAAACGAGCTTAACGGCT 855
QY 839 GAAATTGACCCGCTGGGAGATACGCTCGCAAGATTAACTCAAGAGAAATTGACGGGGAC 898
DB 856 GAAATGACCCGCTGGGAGATACGCTCGCAAGATTAACTCAAGAGAAATTGACGGGGAC 915
QY 899 CCGCACAAGCGGGGATTAATGTGATTAATTCGATGCAACGCGAAGAACTTAACCTTAACCC 958
DB 916 CCGCACAAGCGGGGATTAATGTGATTAATTCGATGCAACGCGAAGAACTTAACCTTAACCC 975
QY 959 TTGACATGTAGCAAAATTTCTAGAGATTAATGATG-CTTCGGGAAACGCTTAACAGAGTG 1017
DB 976 TTGACATGTCTGGAACCTTGTGAGATTAACCTTGCTCCTTCGGGACCGAAGACAGGTG 1035
QY 1018 CTGCATGCTGTGTGTCACTCGCTGTGTGTGAGATGTTGGGTAAAGTCCCGAAGAGCGCA 1077
DB 1036 CTGCATGCTGTGTGTGTCACTCGCTGTGTGTGAGATGTTGGGTAAAGTCCCGAAGAGCGCA 1095
QY 1078 ACCCTGTCACTTAATGGCATCATTTGGTTGGGCACTTAAATGAGACTGCGGTGACAAA 1137
DB 1096 ACCCTGTCACTTAATGGCATCATTTGATTTGGGCACTTAAATGAGACTGCGGTGACAAA 1155
QY 1138 CCGAAGGAAGTGGGATGACGTCAAGTCTCATATGACCTTAATGAGGCTTCAACAGC 1197
DB 1156 CCGAAGGAAGTGGGATGACGTCAAGTCTCATATGACCTTAATGAGGCTTCAACAGC 1215
QY 1198 TAAATCAATGGCGCTACAGAGGTTGCCAACCCCGAGAGGGGAGCTTAATCTCAAGAAAGC 1257
DB 1216 TCATACAAATGGTGGTACAGAGGTTGCCAACCCCGAGAGGTGAGCCCAATCCAGAAAGC 1275
QY 1258 GCGTCTGATGTCGGATTCGAGTCTGCAACTCGACTCCGTAAGTCCGATCGTAGTAAT 1317
DB 1276 GATCTGTAATGTCGGATTTGAGTCTGCAACTCGACTCCATTAAGTCCGATCGTAGTAAT 1335
QY 1318 CGCGATCAGCATGTCGCGGTGAATAGCTTCCGGGTCTTGTACACACCGCGCTCAC 1377
DB 1336 CGCGATCAGCATGTCGCGGTGAATAGCTTCCGGGTCTTGTACACACCGCGCTCAC 1395
QY 1378 CATGGAGTGGGTTTCAACGAAAGCAGTGTCTTAACCGTAAGAGAGGGCGCTTCCACGG 1437
DB 1396 CATGGAGTGGGTTTCAACGAAAGTGTGTGTAACTTCGGAGGGCGCTTACACAGG 1455
QY 1438 TGAGATTCAATGCTGGGGTG 1457
DB 1456 TGAGATTCAATGCTGGGGTG 1475

```

#### RESULT 14

ID AAC86026 standard; cDNA; 1460 BP.

AC AAC86026;

DT 06-AUG-2003 (revised)

DT 29-AUG-2001 (first entry)

DE 16S rDNA, SBR220.

KW 16S rDNA; polyphosphate accumulating organism; PAO; probe; primer;

OS detection; phosphorus; waste water; sludge; ss.

PN unidentified.

PD WO200146459-A1.

PF 28-DEC-2000; 2000WO-AU001611.

PR 23-DEC-1999; 99AU-00004867.

PA (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.

PI Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ;  
 XX DR WPI; 2001-408656/43.  
 XX PT Novel oligonucleotide probe or primer useful for detecting polyphosphate  
 PT accumulating organism in a sample, comprises a sequence that is unique to  
 PT 16S rDNA of polyphosphate accumulating organisms.  
 XX PS Claim 4; Fig 3; 54pp; English.  
 CC The sequences given in AAc86021-30 represent 16S rDNA sequences from  
 CC polyphosphate accumulating organisms (PAOs). Sequences which are unique  
 CC to these 16S rDNA sequences are used to create a probe or primer for  
 CC detecting the relevant organisms. The primer/probe sequences are useful  
 CC for detecting PAO cells in a sample, by treating cells in the sample to  
 CC fix cellular contents, contacting fixed cells with the primer/probe which  
 CC is labelled with a radiolabel, a reporter group or a hapten, under  
 CC conditions which allow the probe to hybridize with 16S rRNA within the  
 CC fixed cell, removing unhybridized probe from the fixed cells, and  
 CC detecting the labeled probe-RNA hybrid by fluorescence in situ  
 CC hybridization. The primer/probe sequences are useful for identifying PAOs  
 CC that are capable of biologically removing phosphorus from waste water.  
 CC Rapid assessment of the presence of a number of PAOs in a waste water  
 CC sample, can be done using the primer/ probe sequences. They allow quick  
 CC and convenient assessment of whether a sludge or waste water sample  
 CC includes PAOs and allows quantitation of PAO cells in samples. (Updated  
 CC on 06-AUG-2003 to correct OS field.)  
 CC XX  
 SQ Sequence 1460 BP; 359 A; 338 C; 475 G; 288 T; 0 U; 0 Other;  
 Query Match 81.7%; Score 1190.8; DB 4; Length 1460;  
 Best Local Similarity 89.8%; Pred. No. 0;  
 Matches 1313; Conservative 0; Mismatches 142; Indels 7; Gaps 3;  
 QY 1 ATTGAACGCTGGCGGCGCTTTACATCAAGTGAACGCGACGACGGATGCTTCAT 60  
 DB 1 ATTAAAGCTGGCGGCGCTTTACATCAAGTGAACGCGACGCGGCGCAACCC-- 58  
 QY 61 CTGTGGCGAGTGGCGGCGGAGTGAATGATCGGAAGTATCCGAAGAGAGGGGGA 120  
 DB 59 CTGTGGCGAGTGGCGGCGGAGTGAATGATCGGAAGTATCCGAAGTATCGGAGTA 118  
 QY 121 ACCGATGGAAGATGCTCTAATACCGCATATCTTAAGAGAGAAAGCAGGGATCGAA 180  
 DB 119 ACCGAGGAAGATGCTCTAATACCGCATATCTTAAGAGAGAAAGCAGGGATCGAA 178  
 QY 181 GACCTTGCGCTTTTGAAGCGGCGGATGCTGATTAAGTGAAGTGGGATTAAGGCTTAC 240  
 DB 179 GACCTTGCGCTTTTGAAGCGGCGGATGCTGATTAAGTGAAGTGGGATTAAGGCTTAC 238  
 QY 241 CAAGGCGAGCATGATGATGTTGTTCTGAGAGAGACGACCAACTGGGATCTGAGACAG 300  
 DB 239 CAAGGCGAGCATGATGATGTTGTTCTGAGAGAGATCCGACACTGGGATCTGAGACAG 298  
 QY 301 GCCAGACTCTCTACGAGGAGCAGACAGTGGGGAATTTTGGACAATGGGCGGAACCTGATC 360  
 DB 299 GCCAGACTCTCTACGAGGAGCAGACAGTGGGGAATTTTGGACAATGGGCGGAACCTGATC 358  
 QY 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCGAGTTGTAAGCTCTTTCAGTGAAGAG 420  
 DB 359 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCGAGTTGTAAGCTCTTTCAGTGAAGAG 418  
 QY 421 AAAGTTACGGTTAAATATGTAACCATGACGGTATCGACAGAAAGACCGGCTAAC 480  
 DB 419 AATGTGTTGGTTAAATACCTGATGATGACGGTATCGACAGAAAGACCGGCTAAC 478  
 QY 481 TACGTGCAAGCAGCGCGGTATATCGAGGGTCAAGCGTTATCGAAATTAATCTGGGCGT 540  
 DB 479 TACGTGCAAGCAGCGCGGTATATCGAGGGTCAAGCGTTATCGAAATTAATCTGGGCGT 538  
 QY 541 AAAGGTTGCGCAGCGCGCTTGTGAATGATGATGTAATCCCGGGCTTAACCTGGGAAT 600  
 DB 539 AAAGGTTGCGCAGCGCGCTTGTGAATGATGATGTAATCCCGGGCTTAACCTGGGAAT 598

QY 601 TGGCTTGAACCTCAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTTAGCAG 660  
 DB 599 TGCATTTGAGACTGCAAGACTGAGATTTTGGCAGAGGAGGAGTGAATTCACCTGTAGCAG 658  
 QY 661 TGAATGCTAGAGATATGAGAAACATGATGTCGAGCAGACCTCTCGGTTAACT 720  
 DB 659 TGAATGCTAGAGATATGAGAAACATGATGTCGAGCAGACCTCTCGGTTAACT 718  
 QY 721 GACGCTCATGCAAGAAAGCTGGGAGCAAAAGATTAAGTATCCTGTTAGTCCAGCGC 780  
 DB 719 GACGCTCATGCAAGAAAGCTGGGAGCAAAAGATTAAGTATCCTGTTAGTCCAGCGC 778  
 QY 781 CTAAAGATGTAACATGATTTGG--GCTTATTAAGCTTGGTAAACGAAGTAAACGGCT 838  
 DB 779 CTAAAGATGTAACATGATTTGGAGGAGGTTAACTTTTATGTCGTTAGTAAACGGCT 838  
 QY 839 GAAGTTGACCGGCTGGGAGTACGGTCCGAAATTAACCTAAAGAAATTAAGCAGGAGC 898  
 DB 839 GAAGTTGACCGGCTGGGAGTACGGTCCGAAAGTAACTAAAGAAATTAAGCAGGAGC 898  
 QY 899 CCGCACAAGCGGTGATTAATGATTAATTCGATGCAACGCGAAACCTTAACCTACC 958  
 DB 899 CCGCACAAGCGGTGATTAATGATTAATTCGATGCAACGCGAAACCTTAACCTACC 958  
 QY 959 TTGACATGTAAGGAATTTCTAGATATGATTAATTTGCT--TCGGGAACGTTAACAGG 1015  
 DB 959 TTGACATGTAAGGAATTTCTAGATATGATTAATTTGCT--TCGGGAACGTTAACAGG 1018  
 QY 1016 TGTGATGATGCTGCTGAGCTGCTGCTGAGATGTTAGTTCGCGCAAGAGC 1075  
 DB 1019 TGTGATGATGCTGCTGAGCTGCTGCTGAGATGTTAGTTCGCGCAAGAGC 1078  
 QY 1076 CAACCTTGTCTTAATTCGATCATTTGTTGGGCACTTTATAGATGCTCGGTGACA 1135  
 DB 1079 CAACCTTGTCTTAATTCGATCATTTGTTGGGCACTTTATAGATGCTCGGTGACA 1138  
 QY 1136 AACCGGAGGAGTGGGAGTGAAGTCAAGTCTCAAGTCCGCTTATGGGAGGCTTCA 1195  
 DB 1139 AACCGGAGGAGTGGGAGTGAAGTCAAGTCTCAAGTCCGCTTATGGGAGGCTTCA 1198  
 QY 1196 CGTAATACATGCGCGGTACAGAGGTTGCCAACCCGCGAGGGAGGATTAATTCAGAA 1255  
 DB 1199 CGTAATACATGCGCGGTACAGAGGTTGCCAACCCGCGAGGGAGGATTAATTCAGAA 1258  
 QY 1256 GCGCGTGTAGTCCGATGCGAGTGTGCAACTGCACTCCGTAAGTCGAATCGCTAGTA 1315  
 DB 1259 GCGCGTGTAGTCCGATGCGAGTGTGCAACTGCACTCCGTAAGTCGAATCGCTAGTA 1318  
 QY 1316 ATCGGGATCAGCATGTCGCGGTGAATACGTTCCGCGGCTTTGACACACCGCCGCTAC 1375  
 DB 1319 ATCGGGATCAGCATGTCGCGGTGAATACGTTCCGCGGCTTTGACACACCGCCGCTAC 1378  
 QY 1376 ACCATGAGAGTGGTTTCAACGAAGCAGTATTAACCGTAAGAGAGGCGCTTGCCAC 1435  
 DB 1379 ACCATGAGAGTGGTTTCAACGAAGCAGTATTAACCGTAAGAGAGGCGATTAACAC 1438  
 QY 1436 GGTGATTCATGACTGGGGTG 1457  
 DB 1439 GGCAGGGTTCTGACTGGGGTG 1460  
 RESULT 15  
 ABL40355  
 ID ABL40355 standard; DNA; 1464 BP.  
 XX ABL40355;  
 AC ABL40355;  
 DT 01-JUL-2002 (first entry)  
 DX Sequence related to an oligonucleotide for the detection of *Azarcus* sp.  
 XX Selenic acid; Ma-23; ds.  
 KW

XX Azarcus sp.  
OS JP2001346584-A.  
XX 18-DEC-2001.  
XX 07-JUN-2000; 2000JP-00170641.  
XX 07-JUN-2000; 2000JP-00170641.  
XX (MITO) MITSUBISHI JUKOGYO KK.  
XX WPI; 2002-199319/26.  
XX Novel oligonucleotide useful for specific detection of a selenic acid  
PT reducing microbe.  
XX Example 1; Page 6; 7pp; Japanese.  
XX The invention relates to an oligonucleotide that may be used for  
CC detecting Azarcus sp. strain MA-23, a selenic acid reducing microbe. The  
CC current sequence represents a sequence related to an oligonucleotide for  
CC the detection of Azarcus sp  
XX  
SQ Sequence 1464 BP; 364 A; 338 C; 470 G; 292 T; 0 U; 0 Other;

Query Match 81.4%; Score 1186; DB 6; Length 1464;  
Beeb Local Similarity 90.0%; Pred. No. 0; Mismatches 140; Indels 7; Gaps 4;  
Matches 1317; Conservative 0;

Qy 1 ATTGAACGCTGGCGGCAATGCTTTACACATGCAAGTCGAAACGGACGACGATGCTTGAT 60  
Db 1 ATTGAACGCTGGCGGCAATGCTTTACACATGCAAGTCGAAACGGACGACGATGCTTGAT 60  
Qy 61 CTGCTGGCGAGTGGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 61 CTGCTGGCGAGTGGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Qy 121 ACGCATGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
Db 121 ACGTACGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
Qy 181 GACCTTGGCTTTTGGAGCGGCGGATGCTGATGATGATGATGATGATGATGATGATGAT 240  
Db 181 GACCTTGGCTTTTGGAGCGGCGGATGCTGATGATGATGATGATGATGATGATGATGAT 240  
Qy 241 CAAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
Db 241 CAAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
Qy 301 GCCGCAATCTCTTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
Db 301 GCCGCAATCTCTTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
Qy 361 CAGCAATGCGGAGTGAAGAAAGGCTTGGGTTGTAAGCTTTTCACTGAGGAAGA 420  
Db 361 CAGCAATGCGGAGTGAAGAAAGGCTTGGGTTGTAAGCTTTTCACTGAGGAAGA 420  
Qy 421 AAAGGTTACGTTAAATTCGTGACCATGACGATGACGATGACGATGACGATGACGAT 480  
Db 421 AAAGGTTACGTTAAATTCGTGACCATGACGATGACGATGACGATGACGATGACGAT 480  
Qy 481 TACGTCGCAAGGCGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
Db 481 TACGTCGCAAGGCGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
Qy 541 AAAGGTTGCGAGCGGCTTGTAGTGAATGATGATGATGATGATGATGATGATGATGAT 599  
Db 541 AAAGGTTGCGAGCGGCTTGTAGTGAATGATGATGATGATGATGATGATGATGATGAT 600  
Qy 600 TTGCGTTTGAATCAAAAGTTAGT GTGCGAGGAGGAGGATGATGATGATGATGATGAT 658  
Db 600 TTGCGTTTGAATCAAAAGTTAGT GTGCGAGGAGGAGGATGATGATGATGATGATGAT 658

Db 601 TTGCGTTTGGACTGCGAGGCTTGAAGTACGCGCAGAGGCGGATGATGATGATGATGAT 660  
Qy 659 AGTGAATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
Db 661 AGTGAATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Qy 719 CTGACGCTCAATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 778  
Db 721 CTGACGCTCAATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
Qy 779 CCTTAAACGATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 836  
Db 781 CCTTAAACGATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
Qy 837 GTGAAGTTGACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 896  
Db 841 GTGAAGTTGACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
Qy 897 ACCGCGACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 956  
Db 901 ACCGCGACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Qy 957 CTTGACATGATGAGGATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGAT 1013  
Db 961 CTTGACATGATGAGGATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGAT 1020  
Qy 1014 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073  
Db 1021 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
Qy 1074 CGCAACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133  
Db 1081 CGCAACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Qy 1134 CAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1193  
Db 1141 CAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
Qy 1194 CAGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253  
Db 1201 CAGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
Qy 1254 AAGGCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313  
Db 1261 AAGGCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
Qy 1314 TATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373  
Db 1321 TATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
Qy 1374 ACACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1433  
Db 1381 ACACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
Qy 1434 ACGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
Db 1441 ACGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1464

RESULT 16  
AAC86028  
ID AAC86028 standard; cDNA; 1459 BP.  
XX AAC86028;  
AC 06-AUG-2003 (revised)  
DT 29-AUG-2001 (first entry)  
XX 16s rDNA, SBRB34.  
XX 16s rDNA; polyphosphate accumulating organism; PNO; probe; primer;  
KM detection; phosphorus; waste water; sludge; ss.

OS Unidentified.  
XX XX MO200146459-A1.  
XX XX 28-JUN-2001.  
XX XX 28-DEC-2000; 2000MO-AU001611.  
XX PF 23-DEC-1999; 99AU-00004867.  
XX PR (CRCM-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.  
XX PA Hugenholz P, Crocetti GR, Tyson GW, Blackall LL;  
XX DR WPI; 2001-408656/43.  
XX PT Novel oligonucleotide probe or primer useful for detecting polyphosphate  
XX PT accumulating organism in a sample, comprises a sequence that is unique to  
XX PT 16S rDNA of polyphosphate accumulating organisms.  
PS Claim 4; Fig 3; 54pp; English.  
XX CC The sequences given in AAC86021-30 represent 16S rDNA sequences from  
XX CC polyphosphate accumulating organisms (PAOs). Sequences which are unique  
XX CC to these 16S rDNA sequences are used to create a probe or primer for  
XX CC detecting the relevant organisms. The primer/probe sequences are useful  
XX CC for detecting PAO cells in a sample, by treating cells in the sample to  
XX CC fix cellular contents, contacting fixed cells with the primer/probe which  
XX CC is labelled with a radiolabel, a reporter group or a hapten, under  
XX CC conditions which allow the probe to hybridize with 16S rRNA within the  
XX CC fixed cell, removing unhybridized probe from the fixed cells, and  
XX CC detecting the labeled probe-RNA hybrid by fluorescence in situ  
XX CC hybridization. The primer/probe sequences are useful for identifying PAOs  
XX CC that are capable of biologically removing phosphorus from waste water.  
XX CC Rapid assessment of the presence of a number of PAOs in a waste water  
XX CC sample, can be done using the primer/ probe sequences. They allow quick  
XX CC and convenient assessment of whether a sludge or waste water sample  
XX CC includes PAOs and allows quantification of PAO cells in samples. (Updated  
XX CC on 06-AUG-2003 to correct OS field.)  
SQ Sequence 1459 BP; 360 A; 338 C; 472 G; 288 T; 0 U; 1 Other;  
Query Match 80.8%; Score 1177.8; DB 4; Length 1459;  
Best Local Similarity 89.6%; Pred. No. 0;  
Matches 1300; Conservative 1; Mismatches 143; Indels 7; Gaps 3;  
QY 12 GCGGCATGCTTTACACATGCAAGTCGAAACGCGACACAGATGCTTGATCTGTCGAG 71  
DB 11 GCGGCATGCTTTACACATGCAAGTCGAAACGCGACACAGATGCTTGATCTGTCGAG 68  
QY 72 TGGCGGACGGGTGATGATATCATTCGAAAGCTATCCAGAAAGAGGGGGTAAACGATCGAAA 131  
DB 69 TGGCGGACGGGTGATGATGAAACATCGGAAACGATATCTGAGATGGGGGATTAACGTAAGGAAA 128  
QY 132 GATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAGAACCCTTGCGCT 191  
DB 129 GTTACGCTAATACCGCATATCTCTGAGACGAGAAAGCAGGGGATCGAAGAACCCTTGCGCT 188  
QY 192 TTTGAGCGCGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGGCTTACCAAGCGGACGA 251  
DB 189 CTGGAGCGCGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGGCTTACCAAGCGGACGA 248  
QY 252 TCAGTATGTTGGTGTGAGAGGACGACCACTGGGAATTGAGACACGGGCCCAAGTCTC 311  
DB 249 TCCGTATGCGGGTCTGAGAGGATATCCGCACACTGGGACTGAGACACGGGCCCAAGTCTC 308  
QY 312 TACGGAGGACGAGAGGGGAAATTTGGAACAATGGGGCGCAAGCTGATCCAGCAATGGCG 371  
DB 309 TACGGAGGACGAGAGGGGAAATTTGGAACAATGGGGCGCAAGCTGATCCAGCAATGGCG 368  
QY 372 CGTAGTGAAGAAGGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAAAGAAAGTTACGG 431  
DB 369 CGTAGTGAAGAAGGCTTCGGGTTGTAAAGCTCTTTCGACGGGAAAGAAATTCGACGGG 428

QY 432 TAAATATGTCGACCCCATGACGGTATCGACAGAAAGACCGGCTTAACCTAGTGCACG 491  
DB 429 TTATATCCTGTGTAGATACCGATACCGATTAAGAAAGACCGGCTTAACCTAGTGCACG 488  
QY 492 AGCCGCGTAAATACGTAAGGATGCAAGCGTTAAATCGGAATTAATCTAGGCGTAAAGGATGCC 551  
DB 489 AGCCGCGTAAATACGTAAGGATGCAAGCGTTAAATCGGAATTAATCTAGGCGTAAAGGATGCC 548  
QY 552 AGCCGCGTAAATACGTAAGGATGCAAGCGTTAAATCGGAATTAATCTAGGCGTAAAGGATGCC 611  
DB 549 AGCCGCGTAAATACGTAAGGATGCAAGCGTTAAATCGGAATTAATCTAGGCGTAAAGGATGCC 608  
QY 612 CTACAAAGCTAAGTGTGACAGAGGAGGTGAAATTCATGTGTAGCAGTGAATATCGTA 671  
DB 609 CTGCGACGCTGAGATTTGGCAGAGGGGGGTGAAATTCACGTGTAGCAGTGAATATCGTA 668  
QY 672 GAGATATGGAAGAACATCGATGCGAAGGACGCTTCCTGGTTAAACATGACGCTCATGC 731  
DB 669 GAGATATGGAAGAACATCGATGCGAAGGACGCTTCCTGGTTAAACATGACGCTCATGC 728  
QY 732 ACGAAAGCTGGGGAGACAAACAGATTAATACCTGTGTATGTCACAGCCCTTAAACGATGT 791  
DB 729 ACGAAAGCTGGGGAGACAAACAGATTAATACCTGTGTATGTCACAGCCCTTAAACGATGT 788  
QY 792 CAATAGTGTGG--GCCTTATAGGCTTGGTAAACGAAAGCTAACGCGTGAAGTTGACCG 849  
DB 789 CAATAGTGTGGTGGAGGGTTAAACCTTTAATGCTGTAGCTAACCGGTGAAGTTGACCG 848  
QY 850 CCTGGGAGTACGCTGCGAAGATTAATACTCAAAGAAATGACGGGACCGGACCAACGGC 909  
DB 849 CCTGGGAGTACGCTGCGAAGATTAATACTCAAAGAAATGACGGGACCGGACCAACGGC 908  
QY 910 GTGATATATGATTAATTCATGACAGCGGAAACCTTACCTAACCTTTGACATGTAG 969  
DB 909 GTGATATATGATTAATTCATGACAGCGGAAACCTTACCTAACCTTTGACATGTAG 968  
QY 970 CGAATTTTATAGATAGATTAATGCT---TCGGGAAAGCTAACACAGGTGCTGATGAC 1026  
DB 969 CGAATTTTATAGATAGATTAATGCT---TCGGGAAAGCTAACACAGGTGCTGATGAC 1028  
QY 1027 TGTCTGACCTGCTGTCTGAGATGTTGGGTTAAAGTCCCGCAACGACGCAACCTTGTCT 1086  
DB 1029 TGTCTGACCTGCTGTCTGAGATGTTGGGTTAAAGTCCCGCAACGACGCAACCTTGTCT 1088  
QY 1087 ATTAATTCATCATTTTGGTGGGCACTTTAATGAGACTGCGGGTGAACAAACCGGAGGAA 1146  
DB 1089 ATTAATTCATCATTTTGGTGGGCACTTTAATGAGACTGCGGGTGAACAAACCGGAGGAA 1148  
QY 1147 GGTGGGATGAGTCAAGTCTCATGAGCCCTTATGGGTGAGGCTTACACGTTAATCAAT 1206  
DB 1149 GGTGGGATGAGTCAAGTCTCATGAGCCCTTATGGGTGAGGCTTACACGTTAATCAAT 1208  
QY 1207 GCGCGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGGCTGTAG 1266  
DB 1209 GGTGCGTCAAGAGGTTGCAACCCGCGAGGGGAGCAATCTCAGAAAGCGGCTGTAG 1268  
QY 1267 TCCGATTCGAGTCTGCAACTGACTCCGTGAAGTGGAAATGCTTAATGCGGATCA 1326  
DB 1269 TCCGATTCGAGTCTGCAACTGACTCCGTGAAGTGGAAATGCTTAATGCGGATCA 1328  
QY 1327 GCATGTCGAGGATAGAGTTCCCGGGTCTTGTACACACGCGCGTACACCACTGAGGAGT 1386  
DB 1329 GCATGTCGAGGATAGAGTTCCCGGGTCTTGTACACACGCGCGTACACCACTGAGGAGT 1388  
QY 1387 GGGTTTCACAGAGGATAGTCTTAACCGTAAGAGAGGCGCTTGGCAAGTGAATTCGA 1446  
DB 1389 GGGTTTCACAGAGGATAGTCTTAACCGTAAGAGAGGCGCTTGGCAAGTGAATTCGA 1448  
QY 1447 TGAATGGGGTGG 1457  
DB 1449 TGAATGGGGTGG 1459

RESULT 17  
AAC86024  
ID AAC86024 standard; cDNA; 1460 BP.  
XX  
XX AAC86024;  
AC  
XX 29-AUG-2001 (first entry)  
XX  
DE R.tennis 16S rDNA #3.  
XX  
XX 16S rDNA; polyphosphate accumulating organism; PAO; probe; primer;  
XX detection; phosphorus; waste water; sludge; ss.  
XX  
XX Rhodocyclus tennis.  
XX  
XX WO200146459-A1.  
XX  
XX 28-JUN-2001.  
XX  
XX 28-DEC-2000; 2000WO-AU001611.  
XX  
XX 23-DEC-1999; 99AU-00004867.  
XX  
XX (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.  
XX  
XX Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ;  
XX WPI; 2001-408656/43.  
XX  
XX Novel oligonucleotide probe or primer useful for detecting polyphosphate  
PT accumulating organism in a sample, comprises a sequence that is unique to  
PT 16S rDNA of polyphosphate accumulating organisms.  
XX  
XX Claim 4; Fig 3; 54bp; English.  
XX  
XX The sequences given in AAC86021-30 represent 16S rDNA sequences from  
CC polyphosphate accumulating organisms (PAOs). Sequences which are unique  
CC to these 16S rDNA sequences are used to create a probe or primer for  
CC detecting the relevant organisms. The primer/probe sequences are useful  
CC for detecting PAO cells in a sample, by treating cells in the sample to  
CC fix cellular contents, contacting fixed cells with the primer/probe which  
CC is labelled with a radiolabel, a reporter group or a hapten, under  
CC conditions which allow the probe to hybridize with 16S rDNA within the  
CC fixed cell, removing unhybridized probe from the fixed cells, and  
CC detecting the labeled probe-RNA hybrid by fluorescence in situ  
CC hybridization. The primer/probe sequences are useful for identifying PAOs  
CC that are capable of biologically removing phosphorus from waste water.  
CC Rapid assessment of the presence of a number of PAOs in a waste water  
CC sample, can be done using the primer/ probe sequences. They allow quick  
CC and convenient assessment of whether a sludge or waste water sample  
CC includes PAOs and allows quantitation of PAO cells in samples  
XX  
XX Sequence 1460 BP; 366 A; 338 C; 471 G; 283 T; 0 U; 2 Other;  
SQ  
Query Match 80.7%; Score 1176; DB 4; Length 1460;  
Best Local Similarity 89.1%; Pred. No. 0;  
Matches 1303; Conservative 0; Mismatches 152; Indels 7; Gaps 3;  
QY 1 ATTGAACGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGAGATGCTTGAT 60  
DB 1 ATTGAACGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGAGATGCTTGAT 58  
QY 61 CTGTGGGAGTGGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 59 CTGTGGGAGTGGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 118  
QY 121 ACGCATCGAAGATGCTAATATACGATATCTCTAAGAGAGAGAGAGAGAGAGAG 180  
DB 119 ACGTACGAAAGTTCACGCTAATACGCTAATCTCTGAGAGAGAGAGAGAGAGAG 178  
QY 181 GACCTTGGCGCTTTTGAAGCGGCGGATGCTGATGATGATGATGATGATGATGAT 240  
DB 181 GACCTTGGCGCTTTTGAAGCGGCGGATGCTGATGATGATGATGATGATGATGAT

DB 179 GACCTTGGCGCTTTTGAAGCGGCGGATGCTGATGATGATGATGATGATGATGAT 238  
QY 241 CAAAGCGAGATCAATGATGTTGTTCTGAGAGAGACGACCGACATCTGGGACTGAGAC 300  
DB 239 CAAAGCGAGATCAATGATGTTGTTCTGAGAGAGATGATGATGATGATGATGATGAT 298  
QY 301 GCCCAGATCTCTACGGGAGGAGCAGATGGGGGAAATTTTGAAGAGAGAGAGAGAGAG 360  
DB 299 GCCCAGATCTCTACGGGAGGAGCAGATGGGGGAAATTTTGAAGAGAGAGAGAGAGAG 358  
QY 361 CAGCAATGCCCGCTGATGAG 420  
DB 359 CAGCCATGCCCGCTGATGAG 418  
QY 421 AAGGTTACGGTAATTAATCTGATCCATGATGATGATGATGATGATGATGATGATGAT 480  
DB 419 AATTTGCTCAGAGATTAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 478  
QY 481 TAGCTGCGACAGACCGCGGATTAATCGTAGGGGTCAGAGCGTTAATCGGATTTACTGG 540  
DB 479 TAGCTGCGACAGACCGCGGATTAATCGTAGGGGTCAGAGCGTTAATCGGATTTACTGG 538  
QY 541 AAGGGTGGCGAGCGGCTTTGTAAGTCAAGATGTAATCCCGGGCTTAACCTGGAGAT 600  
DB 539 AAGCGTGGCGAGCGGCTTTGTAAGTCAAGATGTAATCCCGGGCTTAACCTGGAGAT 598  
QY 601 TGCCTTTGAAACTACAAAGCTAGATGATGATGATGATGATGATGATGATGATGATGAT 660  
DB 599 TGCCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 658  
QY 661 TGAATTCGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
DB 659 TGAATTCGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
QY 721 GAGCTCATGACAGAAAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
DB 719 GAGCTCATGACAGAAAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778  
QY 781 CTAACGATGATCAATGATGTTG- -GCGCTTATTAGGCTTTGTAAGAGAGAGAGAGAG 838  
DB 779 CTAACGATGATCAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 838  
QY 839 GAACTTGAAGCGGCTGGGAG 898  
DB 839 GAACTTGAAGCGGCTGGGAG 898  
QY 899 CCGCAG 958  
DB 899 CCGCAG 958  
QY 959 TTGACATGATGCGAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATG 1015  
DB 959 TTGACATGATGCGAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATG 1018  
QY 1016 TGTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075  
DB 1019 TGTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078  
QY 1076 CAACCTTGTCAATTAATTTGCCATGATTTGTTGGGACTTAATGAGAGAGAGAGAGAGAG 1135  
DB 1079 CAACCTTGTCAATTAATTTGCCATGATTTAGTTGGGACTTAATGAGAGAGAGAGAGAGAG 1138  
QY 1136 AAGCGAG 1195  
DB 1139 AAGCGAG 1198  
QY 1196 GCTAATACATGAGCGGCTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1255  
DB 1199 GCTAATACATGAGCGGCTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1258  
QY 1256 GCGGCTGTATGCTGAG 1315  
DB 1259 GCGGCTGTATGCTGAG 1318

Oy		1316 ATCGGGATCAGCATGTCGCCGTGAATACCTTTCCGGGCTTGTATCAACAACGCCCGCTCAC	1375
Dd		1319 ATCGGGATCAGCATGTCGCCGTGAATACCTTTCCGGGCTTGTATCAACAACGCCCGCTCAC	1378
Oy		1376 ACCATGGAGTGAGGATTTCACAGAAGCAGTAGTAACCATAACCGAAGAGAGGCCTTGCCAC	1435
Dd		1379 ACCATGGAGTGAGGATTTCGCCAAAGTAAGTTAGCTTAACCCGAAAGAGGCGATTAACAC	1438
Oy		1436 GGTGAGATTTCATGACTGAGGGTG 1457     	
Dd		1439 GGcAggGTTgCcTgaCtGggGtg 1460	
RESULT 18			
ID	ADRA45500	standard; DNA; 1530 BP.	
XX	ADRA45500;		
XX	18-NOV-2004	(first entry)	
DE	16S rRNA gene 357F-518r region DNA fragment SeqIDB9.		
KX	357F-518r: 16S rRNA; beta proteobacterium; ammonia oxidising bacteria; activated sludge; ammonia liquid treatment plant; chemical oxygen demand; COD; reduction; nitrification; denitrifying; ds.  Unidentified.  JP2004242578-A.  PN JP2004242578-A.  PP 02-SEP-2004.  PF 13-FEB-2003; 2003JP-00035713.  PR 13-FEB-2003; 2003JP-00035713.  PA (YAMA ) NIPPON STEEL CORP.  WPI; 2004-620179/60.		
Pt	Noel DNA fragment of microorganisms existing in activated sludge of ammonia liquid treatment plant, useful as index microorganisms for evaluating nitrification or denitrifying capability of ammonia liquid.		
PS	Claim 45; SEQ ID NO 89; 133bp; Japanese.		
CC	This invention relates to a novel DNA fragment comprising the 357E-518R region of the 16S RNA gene of beta proteobacteria, belonging to the ammonia oxidizing bacteria group, or CFB Bacteroides where bacteria exists in activated sludge of an ammonia liquid treatment plant and used for chemical oxygen demand (COD) reduction. The invention is useful in the identification of microorganisms as nitrification or denitrifying index microorganisms for evaluating the nitrification or denitrifying capability of ammonia liquid of the activated sludge by fluorescence in situ hybridisation (FISH). The invention is also useful for developing apparatus for the processing of ammonia liquid. The DNA fragment enables evaluation of the nitrification or denitrifying capability of microorganisms. The present sequence is that of a 16S rRNA gene 357f-518r region of the invention.		
SQ	Sequence 1530 BP; 406 A; 337 C; 474 G; 313 T; 0 U; 0 Other:		
Query Match	79.6%; Score 1159.4; DB 13; Length 1530;		
Best Local Similarity	91.4%; Pred.No. 0;		
Matches 1331; Conservative	0; Mismatches 111; Indels 15; Gaps 9;		
Oy		11 GCGGCAATGCTTTACATGCAAGTCGAAGCGGACACGATGCTTGCATCTGGTGGCGA	70
Dd		29 GCGGCAATGCTTTACATGCAAGTCGAAGCGGACGCGGG--GCTTAGGCGCTGCGGCGCA	86
Oy		71 GTGGCGGACGGGTGATATCATTCGAAAACCTATCCAGAAGAGGGGGGTAAACGATCGAA	130

Db	87	GTGGGGAACGGGTGTGATATACATGGAACGTGTCTTAAGTGGGGAATTAAGCATCGAA	146
OY	131	AGATGTGCTAATAACCGCAATATACTCTMAAGAGAGAAAGCGGGATTCGAAAGACTTGGCGC	190
Db	147	AGATGTGCTAATAACCGCAATAT-CTTCCAGAGAGAAAGCGGGGATTCGAAAGACTTGGCGC	205
OY	191	TTTTGAGACGGCCGATGTCTGATTAGTAGTTGTGTGGGTAAAGGCTTACCAAGCGACG	250
Db	206	TAAAGAGAGGGCCGATGTCTGATTAGTAGTTGTGTGGGTAAAGGCTTACCAAGCGACG	265
OY	251	ATCAGTAGTTGTCCTGAGAGGACGACCAAGCCACTGGGACTGAGACACGGCCAGACTC	310
Db	266	ATCAGTAGTTGTCCTGAGAGGACGATCAGCCACACTGGGACTGAGACACGGCCAGACTC	325
OY	311	CTACGGGAGGACGAGTGGGGGAAATTTTGGACAAATGGGGCGAAGCCGTATCCAGCAATGCG	370
Db	326	CTACGGGAGGACGAGTGGGGGAAATTTTGGACAAATGGGGCGAAGCCGTATCCAGCCATGCGC	385
OY	371	GGGTGAGTGAAGAAAGGCCCTTCGGGTGTGTAAGCTCTTTCAGTCGAGAAAGAAAAGTTACG	430
Db	386	GGGTGAGTGAAGAAAGGCC-TCGGGTGTGTAAGCTCTTTCAGTTGGAAGAAAGAAATTACG	444
OY	431	GTAATTAATCGTGACCCATGACGGTATGACAGAAAGAACACCGGCTAATCTACGTGCCAG	490
Db	445	GTTAATACTTAATTTATGACGGTATCAACAGAAAAGAACACCGGCTAATCTACGTGCCAG	504
OY	491	CAGCCGGGGTATATAGTAGGGGTGCAAGGGTTAATGCGAATTATCTGGGGGTAAAGGGTGGC	550
Db	505	CAGCCGGGGTATATAGTAGGGGTGCAAGGGTTATGCGAATTATCTGGGGGTAAAGGGTGGC	564
OY	551	CAGCGCGCTTGTAAAGTACAGATGTGTAAATCCCGGGCTTAACTTGGGAATTCGCTTGA	610
Db	565	CAGCGCGCTTGTAAAGTACAGATGTGTAAAGCCCGGGCTTAACTTGGGAATTCGCTTGA	624
OY	611	ACTACAAAGCTAAGAGTGTGGCAGAGGGAGGTGGAATTCATGTGTAGCAGTGAATTCGT	670
Db	625	ACTGCAAGGCTAAGAGTGTGCAACAGAGGGAGTGGAAATTCATGTGTAGCAGTGAATTCGT	684
OY	671	AGAGTATAGGAAGAACATCGATGGCGAAGGGACGCTCTGGGGTTAACTGACGCTCATG	730
Db	685	AGAGTATAGGAAGAACACCGAATGGCGAAGGGACGCTCTGGGGTTAACTGACGCTCATG	744
OY	731	CACGAAAGCGTGGGGAGCAACAGGATTAGATACCTCGTAGTCCACGCCCTTAAACGATG	790
Db	745	CACGAAAGCGTGGGGAGCAACAGGATTAGATACCTCGTAGTCCACGCCCTTAAACGATG	804
OY	791	TCAACTAGTTGTGGGCTTATTA--GGCTTGTATCGAAGCTAACCGGTGAGTTGACC	848
Db	805	TCAACTAGTTGTGGGCTTATTAAGGATTTGGTAACTAGTAACTAACCGGTGAGTTGACC	864
OY	849	GGCTGGGAGATACGCTGGCAAGATTAAACCTCAAGGAATTGACGGGGACCCGCAACAGC	908
Db	865	GGCTGGGAGATACGCTGGCAAGATTAAACCTCAAGGAATTGACGGGGACCCGCAACAGC	924
OY	909	GGTGGATTATGTGATTATTCGATGCAACGCAAGAAAACCTTACCTACCTTGCATGTA	968
Db	925	GGTGGATTATGTGATTATTCGATGCAACGCAAGAAAACCTTACCTACCTTGCATGTC	984
OY	969	GCGAATTTTCTAGAGATAGATTAGTC--TTCCGGGAACGCTAACACAGGTGCTGCATGG	1022
Db	985	TGGAATCTTAATGAGAGACATTAAGAGTGGCCGTAAAGGGAACCGGGACACAGGTGCTGCATGG	1044
OY	1026	C-TGTGCGACGCTGTGTGCGTAGATATGTGGGTTAAAGTCCCGCAACGA--GCCGAACCC	1088
Db	1045	CGTGTGCGACGCTGTGTGCGTAGATATGTGGGTTAAAGTCCCGCAACGAAGCCGAACCTGT	1104
OY	1082	TTGTCAATTAATTTGCATC-ATTGGTGGGCACTTAAATGAGACGTCGGGTGACAAACCG	1144
Db	1105	TGTGTGTTAATTTGCTATCCATTTTAATAGAGACTTTTAACGAGACGTCGGGTGACAAACCG	1166
OY	1141	GAGGAAGGTGGGATGACGTCAAGTCTCATGCGCTTATAGGGGTACAGCTTCAACGTTA	1200

Db 1165 GAGGAAGTGGGATGACGTC-AGTCTCATGSCCTTATGGTAGGGCTTACACGTAA 1223  
Qy 1201 TACATGCGCGCTACAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAGAAAGCGC 1260  
Db 1224 TACATGCGCGCTACAGAGGTTGCCAATCCGAGAGGAGCAATCTCAGAAAGCGC 1283  
Qy 1261 TCGTAGTCCGATCGAGTCTGCAACTGACTCCGTAAGTCCGATTCGTAATCCG 1320  
Db 1284 TCGTAGTCCGATCGAGTCTGCAACTGACTCCGTAAGTCCGATTCGTAATCCG 1343  
Qy 1321 GATTCAGCATGTGCGGTGAATACGTTCCCGGCTTGTACACACGCGCGTACACAT 1380  
Db 1344 GGTACGACATGCGCGGTGAATACGTTCCCGGCTTGTACACACGCGCGTACACAT 1403  
Qy 1381 GGGAGTGGTTCACGAGACAGTAGTCTAACCCTAAGAGAGGCGCTTGCCACGCTGA 1440  
Db 1404 GGGAGTGGTTCACGAGACAGTAGTCTAACCCTAAGAGAGGCGCTTGCCACGCTGA 1463  
Qy 1441 GATTTCATGACTGGGCTG 1457  
Db 1464 GGGTCATGACTGGGCTG 1480

RESULT 19  
AAC86022  
ID AAC86022 standard; cDNA; 1460 BP.  
XX  
AC AAC86022;  
XX  
DT 29-AUG-2001 (first entry)  
XX  
DE R. tennis rdna #2.  
XX  
KM 16S rRNA; polyphosphate accumulating organism; PAO; probe; primer;  
XX detection; phosphorus; waste water; sludge; ss.  
XX Rhodocyclus tennis.  
XX OS  
PN W0200146459-A1.  
XX  
PD 28-JUN-2001.  
XX  
PF 28-DEC-2000; 2000WO-AU001611.  
XX  
PR 23-DEC-1999; 99AU-00004867.  
XX  
PA (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.  
XX  
PI Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ;  
XX WPI; 2001-408656/43.  
DR  
XX Novel oligonucleotide probe or primer useful for detecting polyphosphate  
PT accumulating organism in a sample, comprises a sequence that is unique to  
PT 16S rDNA of polyphosphate accumulating organisms.  
XX  
PS Claim 4; Fig 3; 54pp: English.

The sequences given in AAC86021-30 represent 16S rDNA sequences from  
CC polyphosphate accumulating organisms (PAOs). Sequences which are unique  
CC to these 16S rDNA sequences are used to create a probe or primer for  
CC detecting the relevant organisms. The primer/probe sequences are useful  
CC for detecting PAO cells in a sample, by treating cells in the sample to  
CC fix cellular contents, contacting fixed cells with the primer/probe which  
CC is labeled with a radiolabel, a reporter group or a hapten, under the  
CC conditions which allow the probe to hybridize with 16S rRNA within the  
CC fixed cell, removing unhybridized probe from the fixed cells, and  
CC hybridization. The primer/probe sequences are useful for identifying PAOs  
CC that are capable of biologically removing phosphorus from waste water.  
CC Rapid assessment of the presence of a number of PAOs in a waste water  
CC sample, can be done using the primer/ probe sequences. They allow quick  
CC and convenient assessment of whether a sludge or waste water sample

CC includes PAOs and allows quantitation of PAO cells in samples  
XX  
SQ Sequence 1460 BP; 363 A; 340 C; 471 G; 284 T; 0 U; 2 Other;  
Qy Query Match 79.5%; Score 1158.4; DB 4; Length 1460;  
Best Local Similarity 88.4%; Pred. No. 0;  
Matches 1292; Conservative 0; Mismatches 163; Indels 7; Gaps 3;  
Qy 1 ATTGAACGCTGCGGCGATGCTTTTACACATGCAAGTGAAGACGAGACGGAATCTTGAT 60  
Db 1 ATTGAACGCTGCGGCGATGCTTTTACACATGCAAGTGAAGACGGAATCTTGAT 58  
Qy 61 CTGTGCGGCGAGTGGCGGAGCGGAGTAATGATTCGGAACGTAAGAGAGGAGGATG 120  
Db 59 NTGGCGACGAGTGGCGGAGCGGAGTAATGATTCGGAACGTAAGAGAGGAGGATG 118  
Qy 121 ACCGATCGAAAGATGTGCTAATACCGCATATTACTTAAGAGAGAAAGCAGGGATCGAA 180  
Db 119 ACGTAGGAAAGTTAGCGTAAATACCGCATATTCGTGAGCAGGAAAGCAGGGATCTTAG 178  
Qy 181 GACCTTGCGCTTTTGGAGCGCGCGATGTCTGATTAAGTATGTTGGTGGGCTTAAC 240  
Db 179 GACCTTGCGCTTTTGGAGCGCGCGATGTCTGATTAAGTATGTTGGTGGGCTTAAC 238  
Qy 241 CAAGCGGAGATGATGATGTTGTTGAGAGAGACGACGACGACCTGGAGACAGC 300  
Db 239 CAAGCGGAGATGATGATGTTGTTGAGAGAGATGATGATGATGATGATGATGATGATG 298  
Qy 301 GCCCAGACTCTTAACGAGAGCAGCAGTGGGAAATTTTGAACAATGGCGCAAGCTGATC 360  
Db 299 GCCCAGACTCTTAACGAGAGCAGCAGTGGGAAATTTTGAACAATGGCGCAAGCTGATC 358  
Qy 361 CAGCATGCGCGGTGATGAAAGAGCCTTCGGGTTTAAAGCTTTTCACTGAGAAAGA 420  
Db 359 CAGCATGCGCGGTGATGAAAGAGCCTTCGGGTTTAAAGCTTTTCACTGAGAAAGA 418  
Qy 421 AAAGTTACGTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
Db 419 AAAGTTACGTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 478  
Qy 481 TACGTGCGACGAGCGCGGTAAATACGTAAGGTGCAACGTTAATCGAATTAATCTGGCGGT 540  
Db 479 TACGTGCGACGAGCGCGGTAAATACGTAAGGTGCAACGTTAATCGAATTAATCTGGCGGT 538  
Qy 541 AAAGGTGCGGAGCGCGCTTTGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 539 AAAGGTGCGGAGCGCGCTTTGATGATGATGATGATGATGATGATGATGATGATGATG 598  
Qy 601 TCGGTTGAAACTTAAGCTAAGTGTGCGAGAGGAGGATGATGATGATGATGATGATG 660  
Db 599 TCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 658  
Qy 661 TGAATGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
Db 659 TGAATGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 718  
Qy 719 GAGCGTATGACGAAAGCGTGGGAGCAACAGAAATTAAGTATGATGATGATGATGATG 778  
Db 721 GAGCGTATGACGAAAGCGTGGGAGCAACAGAAATTAAGTATGATGATGATGATGATG 780  
Qy 781 CTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838  
Db 779 CTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838  
Qy 839 GAAAGTTACCGCTCGGAGAGTACCGCGCAAGTTAAACTCAAGAAATGACGCGGAC 898  
Db 839 GAAAGTTACCGCTCGGAGAGTACCGCGCAAGTTAAACTCAAGAAATGACGCGGAC 898  
Qy 899 CCGACAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958  
Db 899 CCGACAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958  
Qy 959 TTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1015  
Db 959 TTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1015



```
Db      959  TTACACATGTCAGGAAATCTTGAAGATTTAGAGAGTGCCTCCGAAAGGAAACCTGAACACAGG 1018
Qy      1016  TGTGTCATGGCTGTCTGCACTGCTGTCGTGAAGATTTGGGTTAAGTCCCGCAACGAGCG 1075
Db      1019  TGTGTCATGGCTGTCTGCACTGCTGTCGTGAAGATTTGGGTTAAGTCCCGCAACGAGCG 1078
Qy      1076  CAACCCCTGTCAATTAATGTCATCATTTGTTGGGACCTTAATAGACATGCCGGTGACA 1135
Db      1079  CAACCCCTGTCAATTAATGTCATCATTTGTTGGGACCTTAATAGACATGCCGGTGACA 1138
Qy      1136  AACCGAGAAAGGTGGGATGAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTCACA 1195
Db      1139  AACCGAGAAAGGTGGGATGAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTCACA 1198
Qy      1196  CGTAAATCAATGGCGCTACAGAGGTTGCCAACCCCGGAGGGGAGACTATCTCAGAAA 1255
Db      1199  CGTAAATCAATGGCGCTACAGAGGTTGCCAACCCCGGAGGGGAGACTATCTCAGAAA 1258
Qy      1256  GCGCGTCTGATGTCGGATCGGATCGCACTGCACTCGGTAAGTGGAAATGCTTAGTA 1315
Db      1259  GCGCGTCTGATGTCGGATCGGATCGCACTGCACTCGGTAAGTGGAAATGCTTAGTA 1318
Qy      1316  ATGCGGATCAGCATGTGCGGTGAATACGTTCCGGGTCTTGTACACACCGCCGTCAC 1375
Db      1319  ATGCGGATCAGCATGTGCGGTGAATACGTTCCGGGTCTTGTACACACCGCCGTCAC 1378
Qy      1376  ACCATGGAGTGGGTTTACACGAGAGAGGTAAGTCTTAACTTAAGAGAGGCGCTTGCCAC 1435
Db      1379  ACCATGGAGTGGGTTTACACGAGAGAGGTAAGTCTTAACTTAAGAGAGGCGCTTGCCAC 1438
Qy      1436  GGTGAATTCATGACTGGGGGTG 1457
Db      1439  GGTGAATTCATGACTGGGGGTG 1460
```

## RESULT 20

AAC86021  
ID AAC86021 standard; cDNA; 1460 BP.

AC AAC86021;

DT 29-AUG-2001 (first entry)

XX R. tenuis rdna #1.

KW 16S rdna; polyphosphate accumulating organism; PAO; probe; primer;  
detection; phosphorus; waste water; sludge; ss.

XX Rhodocyclus tenuis.

OS

PN WO200146459-A1.

PD 28-JUN-2001.

PF 28-DEC-2000; 2000MO-AU001611.

PR 23-DEC-1999; 99AU-00004867.

XX (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.

XX Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ;

XX WPI; 2001-408656/43.

XX Novel oligonucleotide probe or primer useful for detecting polyphosphate  
accumulating organism in a sample, comprises a sequence that is unique to  
16S rdna of polyphosphate accumulating organisms.

XX Claim 4; Fig 3; 54pp; English.

CC The sequences given in AAC86021-30 represent 16S rdna sequences from  
polyphosphate accumulating organisms (PAOs). Sequences which are unique

```
CC to these 16S rdna sequences are used to create a probe or primer for  
CC detecting the relevant organisms. The primer/probe sequences are useful  
CC for detecting PAO cells in a sample, by treating cells in the sample to  
CC fix cellular contents, connecting fixed cells with the primer/probe which  
CC is labelled with a radiolabel, a reporter group or a hapten, under  
CC conditions which allow the probe to hybridize with 16S rdna within the  
CC fixed cell, removing unhybridized probe from the fixed cells, and  
CC detecting the labeled probe-RNA hybrid by fluorescence in situ  
CC hybridization. The primer/probe sequences are useful for identifying PAOs  
CC that are capable of biologically removing phosphorus from waste water.  
CC Rapid assessment of the presence of a number of PAOs in a waste water  
CC sample, can be done using the primer/probe sequences. They allow quick  
CC and convenient assessment of whether a sludge or waste water sample  
CC includes PAOs and allows quantitation of PAO cells in samples  
XX  
SQ Sequence 1460 BP; 362 A; 339 C; 471 G; 264 T; 0 U; 4 Other;  
Query Match 79.4%; Score 1157.4; DB 4; Length 1460;  
Best Local Similarity 88.3%; Pred. No. 0;  
Matches 1291; Conservative 0; Mismatches 164; Indels 7; Gaps 3;  
Qy 1 ATTTAAACGTCGCGGATGCTTTTACATGCAATGCAAGCGACGAGATGCTTGAT 60  
Db 1 ATTTAAACGTCGCGGATGCTTTTACATGCAATGCAAGCGATTAACGCGGAAAC-- 58  
Qy 61 CTGGTGGGAGTGGGAGAGGGGTGATATGATCGGAAACGTAATCAGAGAGGGGGGTA 120  
Db 59 NTGGCGACAGTGGCGAAGGGGTGATATGATCGGAAACGTCGCTTAAGTGGGGGATA 118  
Qy 121 ACGCATCGAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGATCGAAA 180  
Db 119 ACGTACGAAATTAATACGCTAATACCGCATATCTGTAGCAGAGAAAGCAGGGATCTTAG 178  
Qy 181 GACCTTGGCTTTTGGAGGCGCGATGCTGTATTAAGTTAGTTGGGGTTAAAGCCTTAC 240  
Db 179 GACCTTGGCTTTTGGAGGCGCGATGCTGTATTAAGTTAGTTGGGGTTAAAGCCTTAC 238  
Qy 241 CAAGCGACGATCAGTATGCTGAGAGAGAGCAGACAGCCACACTGGGACTGAGACAGC 300  
Db 239 CAAGCGACGATCAGTATGCTGAGAGAGAGTCCGCCACACTGGGACTGAGACAGC 298  
Qy 301 GCCCAGACTCTTACCGGAGAGCAGAGTGGGGAATTTTGGACATAGGGCGCAAGCTGTATC 360  
Db 299 GCCCAGACTCTTACCGGAGAGCAGAGTGGGGAATTTTGGACATAGGGCGCAAGCTGTATC 358  
Qy 361 CAGCAATGCGCGGTGATGAGAGAGGCTTGGGTTGTAAAGCTTTTCACTGAGAGAGA 420  
Db 359 CAGCAATGCGCGGTGATGAGAGAGGCTTGGGTTGTAAAGCTTTTCACTGAGAGAGA 418  
Qy 421 AAAGTTACGTAATTAATTCGTAACCCATGACCGTATGACAGAGAGAGCAACCGGCTAAC 480  
Db 419 AAAGTTACGTAATTAATTCGTAACCCATGACCGTATGACAGAGAGAGCAACCGGCTAAC 478  
Qy 481 TACGTGCCAGACCGCGGTAAATACGTAGGGGTCAAGGCTTAATCGAATTAATCTGGGCT 540  
Db 479 TACGTGCCAGACCGCGGTAAATACGTAGGGGTCAAGGCTTAATCGAATTAATCTGGGCT 538  
Qy 541 AAAGGTGGGAGGCGCGCTTGTAAAGTCAATGTAATTCCTGGGCTTAACTTGGGAAT 600  
Db 539 AAAGGTGGGAGGCGCGCTTGTGTAAAGTCAATGTAATTCCTGGGCTTAACTTGGGAAT 598  
Qy 601 TGGCTTTGAAATTAATCAAGCTTGAAGTGTGAGAGAGAGGTGGAATTCATGATGTAAGAG 660  
Db 599 TGGCTTTGTAATGTAAGCTTGAAGTGTGAGAGAGAGGTGGAATTCATGATGTAAGAG 658  
Qy 661 TGAATTTGGTATGATGTAAGAGATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
Db 659 TGAATTTGGTATGATGTAAGAGATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718  
Qy 721 GACGCTCATGACGAAAGCTGGGGAGCAAAAGAGATTAGATCCCTGATGATCAAGCC 780  
Db 719 GACGCTCATGACGAAAGCTGGGGAGCAAAAGAGATTAGATCCCTGATGATCAAGCC 778
```



Oy		781	CTTAACGATGCACCTAGTGGTTG-- -GGCCATTATAGACTTGGAACGAAGCTAACGCCGT	838
Dd		779	CTTAACGATGCACCTAGTGGTTG-- -GGCCATTATAGACTTGGAACGAAGCTAACGCCGT	838
Oy		839	GAACTTGACCCTCGGGGAGTACCGGTGCGACAAGATTAAACTCAAGGAATTGACGGGAC	898
Dd		839	GAACTTGACCCTCGGGGAGTACCGGTGCGACAAGATTAAACTCAAGGAATTGACGGGAC	898
Oy		899	CCGCAACAAGCGGTGAGATTATGTGATTAATTCGATGCAAGCGCAAAAACTTACCTTACCC	958
Dd		899	CCGCAACAAGCGGTGAGATTATGTGATTAATTCGATGCAAGCGCAAAAACTTACCTTACCC	958
Oy		959	TTCGACATGTACCGAATTTTCTAGAGATAAGATTACTGCG-- -TTGGGAAACGCTAACAGG	1015
Dd		959	TTCGACATGTACCGAATTTTCTAGAGATAAGATTACTGCG-- -TTGGGAAACGCTAACAGG	1015
Oy		1016	TGCTGCATGCGCTGCTGCAGCTCGTGTGCTGAGATGTTGGTTTAAGTCCCGCAACGACG	1075
Dd		1019	TGCTGCATGCGCTGCTGCAGCTCGTGTGCTGAGATGTTGGTTTAAGTCCCGCAACGACG	1078
Oy		1076	CAACCCCTGTCAATTATTTGCCATCATTTTGGTTGGGCACTTTAATGAACTGCGCGTGACA	1135
Dd		1079	CAACCCCTGTCAATTATTTGCCATCATTTTGGTTGGGCACTTTAATGAACTGCGCGTGACA	1138
Oy		1136	AACCGGAGGAAGGTGGGAGTAGACGTCAAGTCTCTATAGGCCCTTTAATGGGTAGGGCTTACA	1195
Dd		1139	AACCGGAGGAAGGTGGGAGTAGACGTCAAGTCTCTATAGGCCCTTTAATGGGTAGGGCTTACA	1198
Oy		1196	CGTAAATACATAGGCGCGTACAGAGGGTTGCCAACCGCGAGGGGAGAGTAATCTCAGAAA	1255
Dd		1199	CGTAAATACATAGGCGCGTACAGAGGGTTGCCAACCGCGAGGGGAGAGTAATCTCAGAAA	1258
Oy		1256	GCGCGCTGTAAGTCCCGATCGAGTGTGCAACTCGACTCCGTGAAGTCGAATGCTTAGTA	1315
Dd		1259	GCGCGCTGTAAGTCCCGATCGAGTGTGCAACTCGACTCCGTGAAGTCGAATGCTTAGTA	1318
Oy		1316	ATCGCGGATCAGCATGTGCGCGTGAATACGTTCCCGGGTCTTTGATCACACCGCCCGTAC	1375
Dd		1319	ATCGCGGATCAGCATGTGCGCGTGAATACGTTCCCGGGTCTTTGATCACACCGCCCGTAC	1378
Oy		1376	ACCATGGGAGNGGGTTTCCACGAGAAGAGTATGCTTAACCGTAAAGAGAGGCGCTTGCCAC	1435
Dd		1379	ACCATGGGAGNGGGTTTCCACGAGAAGAGTATGCTTAACCGTAAAGAGAGGCGCTTGCCAC	1438
Oy		1436	GGTGAATTCATGACTGGGGTG 1457	
Dd		1439	GGCAGGGTTCGTGACTGGGGTG 1460	
<hr/>				
RESULT 21				
AAQ26729				
ID	AAQ26729	standard; cDNA; 1532 BP.		
XX	AAQ26729;			
DT	25-JAN-1993	(first entry)		
DE	16S rRNA gene from Bordetella bronchiseptica.			
KW	rapid; examine; samples; test; identify; detect; seq.			
OS	Bordetella bronchiseptica.			
PN	JP04179480-A.			
PD	26-JUN-1992.			
Pf	09-NOV-1990;	90JP-00304758.		
PR	09-NOV-1990;	90JP-00304758.		
PA	(ZENK-) ZENKOKU NOGIO KYODO KUMITAI REN. (NISE-) NIPPON SEIHUN KK.			

Query Match	Best Local Similarity	Score	DB 2	Length	Matches	Conservative	Mismatches	Indels	Gaps
1	ATTGAACCGTGGCGGCATGCTTATACATGCAATGCAAGTGCAGACGGCAGCAGATGCTTGCAT	78.9%	1149.2	1532	1287	0	163	10	3
29	ATTGAACCGTGGCGGCATGCTTATACATGCAATGCAAGTGCAGACGGCAGCAGATGCTTGCAT	78.2%	1149.2	1532	1287	0	163	10	3
61	CTGTGTGCGAGTGTGCGGCGAGCGGCTGAGTATGTCATCGAACTATCCAGAAAGGCGGAGTAT	78.2%	1149.2	1532	1287	0	163	10	3
87	CTGTGTGCGAGTGTGCGGCGAGCGGCTGAGTATGTCATCGAACTATCCAGAAAGGCGGAGTAT	78.2%	1149.2	1532	1287	0	163	10	3
121	ACGCATCGAAGATGTGCTATATCCGCATATATCTTAAAGAGAAAGACGGGGATTCGAA	78.2%	1149.2	1532	1287	0	163	10	3
147	ACTAGCGCGAAAGCGTGGCTAATATCCGCATATCCGCATATCCGCATATCCGCATATCCGCAT	78.2%	1149.2	1532	1287	0	163	10	3
181	GACCTTGGCGCTTTTGAAGCGCGCGAGTGTCTATGCTTATGCTTATGCTTATGCTTATGCTTAT	78.2%	1149.2	1532	1287	0	163	10	3
207	GACCTTGGCGCTTTTGAAGCGCGCGAGTGTCTATGCTTATGCTTATGCTTATGCTTATGCTTAT	78.2%	1149.2	1532	1287	0	163	10	3
241	CAAGCGAGCATGCTAGTGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	78.2%	1149.2	1532	1287	0	163	10	3
267	CAAGCGAGCATGCTAGTGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	78.2%	1149.2	1532	1287	0	163	10	3
301	GCCGAGACTCTTACCGGAG	78.2%	1149.2	1532	1287	0	163	10	3
327	GCCGAGACTCTTACCGGAG	78.2%	1149.2	1532	1287	0	163	10	3
361	CAGCATGCGCGGTGAGTGAAG	78.2%	1149.2	1532	1287	0	163	10	3
387	CAGCATGCGCGGTGAGTGAAG	78.2%	1149.2	1532	1287	0	163	10	3
421	AAAGGTATCGGTAAAT	78.2%	1149.2	1532	1287	0	163	10	3
447	AAAGGTATCGGTAAAT	78.2%	1149.2	1532	1287	0	163	10	3
481	TACGTGCGAGCAGCGCGGTAAAT	78.2%	1149.2	1532	1287	0	163	10	3
507	TACGTGCGAGCAGCGCGGTAAAT	78.2%	1149.2	1532	1287	0	163	10	3
541	AAAGGTGCGCGAGCGCGCTTGTAT	78.2%	1149.2	1532	1287	0	163	10	3
567	AAAGGTGCGCGAGCGCGCTTGTAT	78.2%	1149.2	1532	1287	0	163	10	3
601	TGCGTTTGAAT	78.2%	1149.2	1532	1287	0	163	10	3
627	TGCGTTTGAAT	78.2%	1149.2	1532	1287	0	163	10	3
661	TGAATATGCTAGAT	78.2%	1149.2	1532	1287	0	163	10	3
687	TGAATATGCTAGAT	78.2%	1149.2	1532	1287	0	163	10	3
721	GAGCGTATGCAAGAAAGCGTGGGAG	78.2%	1149.2	1532	1287	0	163	10	3
747	GAGCGTATGCAAGAAAGCGTGGGAG	78.2%	1149.2	1532	1287	0	163	10	3
781	CTAAGAGATGCAATATGTTGTGGCGCTTATATATATATATATATATATATATATATATATATAT	78.2%	1149.2	1532	1287	0	163	10	3

Db	807	CTAAACGATGTAACCTAGCTGTTGGGCGCTTCGCGCCCTTGGTAGGCGACCTAAACGCGTGA	866
Qy	841	AGTTGACCGCCCTGGGGAGTACGGTGCAGAAATTAAACTCAAGAAATTGACGGGGACC	900
Db	867	AGTTGACCGCCCTGGGGAGTACGGTGCAGAAATTAAACTCAAGAAATTGACGGGGACC	926
Qy	901	GCACAAAGCGGTGATTAATGTGGATTAAATTGCATGCACACGGAAAACTTAACCTT	960
Db	927	GCACAAAGCGGTGATTAATGTGGATTAAATTGCATGCACACGGAAAACTTAACCTT	986
Qy	961	GACATGTACCAATTTTCTAGAGATAGATTAGTCT---TCGGGAACGCTTAACACAGGTG	1017
Db	987	GACATGTCTGGAATCCCGAAGAGATTTGGGAGTGTCTCCGACAGAGAACCGGAACACAGGTG	1046
Qy	1018	CTGCATGGCTGTCGTCAAGCTGTGTCGTAGAGATGTTGGATTAAATGTCGCCGACAGAGCGCA	1077
Db	1047	CTGCATGGCTGTGTCGAGCTGTGTCGTAGAGATGTTGGATTAAATGTCGCCGACAGAGCGCA	1106
Qy	1078	ACCCCTGTCAATTAATTTGGCATTATTTGGTTGGGCACTTTAAATGAGACTGCCGATGACAA	1137
Db	1107	ACCCCTGTCAATTAATTTGGCATTATTTGGTTGGGCACTTTAAATGAGACTGCCGATGACAA	1161
Qy	1138	CCGGAGGAAGGTGGGGATGACGTCAAGTCTCTATGCGCCCTTAATGGGTAGGGCTTCAACG	1197
Db	1162	CCGGAGGAAGGTGGGGATGACGTCAAGTCTCTATGCGCCCTTAATGGGTAGGGCTTCAACG	1221
Qy	1198	TAAATCAATAGGCGCGCTACAGAGGGTGTGCGAACCCGCGAGGGGGAGCTAATCTCAGAAAGC	1257
Db	1222	TCATACAAATGCTCGGGAACAGAGGGTGTGCGAACCCGCGAGGGGGAGCTAATCTCAGAAAGC	1281
Qy	1258	GCGTGTAGTCCGGAATCGGAAGTCTGCAACTCGACTCCGTGAAGTCCGAATCGTAGTAAT	1317
Db	1282	CGATGTAGTCCGGAATCGGAAGTCTGCAACTCGACTCCGTGAAGTCCGAATCGTAGTAAT	1341
Qy	1318	CGCGGATCAGCAATGTCGCGGTGAATACGTTCCCGGGTCTTTGTAACAACCGCCGTCACAC	1377
Db	1342	CGCGGATCAGCAATGTCGCGGTGAATACGTTCCCGGGTCTTTGTAACAACCGCCGTCACAC	1401
Qy	1378	CATGGGAATGGGTTTCAACAGAGGATGATCTTAACGTTAAGAGAGGGCGCTTCCACCG	1437
Db	1402	CATGGGAATGGGTTTTCACAGAGGATGATGACTTAACCGCAAGGGGGCGATTAACACCG	1461
Qy	1438	TGAGATTCAATGACTGGGGTG 1457	
Db	1462	TAGATTCAATGACTGGGGTG 1481	
RESULT 22			
ADB61686			
ID	ADB61686 standard; DNA; 1532 BP.		
AC	ADB61686;		
XX			
DT	04-DEC-2003 (first entry)		
XX			
DE	16S rRNA of Bordetella bronchiseptica DNA sequence.		
XX			
KM	enriching mRNA; high quality bacterial mRNA; bacterial gene expression;		
KM	poly-A tail; mRNA purification; oligo-dT capture;		
KM	prokaryote mRNA purification; bridging oligonucleotide; targeting region;		
KM	capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;		
KM	eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;		
KM	28S eukaryotic rRNA bridging oligonucleotide; ds.		
XX			
OS	Bordetella bronchiseptica.		
XX			
PN	MO2003054162-A2.		
XX			
PD	03-JUL-2003.		
XX			
PF	19-DEC-2002; 2002WO-US041014.		
XX			

PR	20-DEC-2001; 2001US-00029397.
XX	(AMBI-) AMBION INC.
PA	Murphy GL, Whitley JP;
PI	WPI, 2003-663255/62.
DR	
XX	
XX	Depleting or isolating targeted nucleic acids e.g. RNA, involves using a
PT	bridging oligonucleotide comprising bridging region and a targeting
PT	region complementary to a targeted nucleic acid, and a capture
PT	oligonucleotide.
XX	
XX	
PS	Claim 4; Page 171-172; 208pp; English.
CC	This invention relates to a novel method for isolating, depleting or
CC	separating a targeted nucleic acid, such as RNA, from a sample
CC	comprising targeted and non-targeted nucleic acids. It effects a way of
CC	enriching for non-targeted nucleic acids such as mRNAs. Isolating
CC	sufficient quantities of high quality bacterial mRNA is a demanding
CC	process which impedes analysis of bacterial gene expression in the
CC	presence of host cells. A small percentage of bacterial mRNAs may be poly
CC	-A tailed, but these are targeted for degradation and tend to be
CC	unstable. As a result, the commonly employed method for mRNA purification
CC	with eukaryotic cells, oligo-dt capture, is ineffective. The present
CC	invention provides an alternative, more suitable method for mRNA
CC	purification from prokaryotes. The method of the invention comprises the
CC	incubation of a sample with a bridging oligonucleotide (containing a
CC	targeting region) and subsequently incubating with a capture
CC	oligonucleotide allowing the isolation of the target from the sample. The
CC	method is useful for depleting or isolating targeted nucleic acid, for
CC	example RNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S
CC	or 18S or eukaryotic 28S RNA, from a sample. The RNA sequence may
CC	comprise any one of 64 fully defined sequences as given in the
CC	specification. The present sequence is that of a DNA sequence which
CC	represents the sequence of 16S RNA of Bordetella bronchiseptica related
CC	to the invention.
XX	
XX	
SQ	Sequence 1532 BP; 378 A; 363 C; 490 G; 301 T; 0 U; 0 Other;
Query Match	78.9%; Score 1149.2; DB 10; Length 1532;
Best Local Similarity	88.2%; Pred. No. 0;
Matches 1287; Conservative	0; Mismatches 163; Indels 10; Gaps 3
OY	1 ATTGAACCTGGCGGCATGCTTTACATCAATGCAAGTCGAAACGGCAGACGATGTTCGAT 60
DB	29 ATTGAACCTGGCGGCATGCTTTACATCAATGCAAGTCGAAACGGCAGACGATGTTCGAT 86
OY	61 CTGTGTGGAGTGGCGGACGGGTGATGATGATCGGAACGTATCCAGAAGAGGGGGATA 120
DB	87 CTGTGTGGAGTGGCGGACGGGTGATGATGATCGGAACGTATCCAGAAGAGGGGGATA 146
OY	121 ACGCATGAAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAACGAGGGGATCGAAA 180
OY	147 ACTACGCCAAAGCGTGCTAATACCGCATATCGCCCTTAAGCGGGGAAACGGGGGACCTTCG 206
DB	181 GACCTTGGGCTTTTGGAGCGGCGCATATGCTGATTAGCTAGTGTGGGGTTAAAGGCTTAC 240
DB	207 GACCTTGGGCTTTTGGAGCGGCGCATATGCTGATTAGCTAGTGTGGGGTTAAAGGCTTAC 266
OY	241 CAAGCGCAGATCAAGTAGTTGGTCTGAGAGAGCACACGACCACACTGGGACTTGAGACAG 300
DB	267 CAAGCGCAGATCCGTACCTGGTTTGAAGAGAGCACACGACCACACTGGGACTTGAGACAG 326
OY	301 GCCCAGACTCTTACCGGAGGCGACAGTGGGGAATTTTGGACAATGGGCGCAACGCTGATC 360
DB	327 GCCCAGACTCTTACCGGAGGCGACAGTGGGGAATTTTGGACAATGGGCGCAACGCTGATC 386
OY	361 CAGCAATCCCGCTGAGAGTGAAGGAAGGCGCTTGGGGTTTGAAGCTCTTTCAGTCGAGAA 420
DB	387 CAGCCATCCCGCTGTGCGATGAGGCGCTTGGGGTTTGAAGCACTTTTGGCAGAGAA 446
OY	421 AAAGGTTACGGTTAATATCTGACCCCATGACGATATGACAGAAAGACCGGCTAAC 480

```
Db      447 AACGGACCGGGCTTAATATCTCTGTCAACTGACGGTACTCTGCAGAAATTAAGACCGGGCTAAC
      481 TAGTGCCACACACCGCGGTAAATACGTAGGGGTGCAACGTTAATCGGAATTACTGGGGCT 540
      507 TAGGTGCACACACCGCGGTAAATACGTAGGGGTGCAACGTTAATCGGAATTACTGGGGCT 566
      541 AAAGGTCGCGACGCGCTTGTAAAGTCAGATGTAATCCCGGGCTTAACTGGGAAT 600
      567 AAAGCGGCCACGCGGTTCGGAAGAAGATGTAATCCAGGGCTTAACTGGGAAC 626
      601 TGGCTTTGAAACTACAAAGCTAGAGTGTGCAGAGGAGGTGCAATTCATGTGTACAG 660
      627 TGCATTTTAACTACCGGGCTAGAGTGTGTAGAGGAGGTGGAATTCGCCGTGTACAG 686
      661 TGAATGCCGTAGATATGGAAGAACATGCATGGCGAAGGACGCTCGGGGTAACT 720
      687 TGAATGCCGTAGATATGCGAGAACACCGATGGCAAGGACGCTCGGGATTAACCT 746
      721 GACGCTCATGCAAGAAAGCGTGGAGCAAAACAGATTAAGTACCTGTAGTCAAGCC 780
      747 GACGCTCATGCAAGAAAGCGTGGAGCAAAACAGATTAAGTACCTGTAGTCAAGCC 806
      781 CTAAACGATGTCAACTAGTTGTTGGCCCTTAATAGCTGTGAACGAAGCTAACGCGTGA 840
      807 CTAAACGATGTCAACTAGTTGTTGGCCCTTAATAGCTGTGAACGAAGCTAACGCGTGA 866
      841 AGTTGACCGCTGGGGAGTACCGGTGCAAGTTAAACTCAAAAGAAATTAAGCGGGACCC 900
      867 AGTTGACCGCTGGGGAGTACCGGTGCAAGTTAAACTCAAAAGAAATTAAGCGGGACCC 926
      901 GCACACGCGGTGATTAATGTGATTAATCGATGCAACGCGAATAAATCTTAACCTT 960
      927 GCACACGCGGTGATTAATGTGATTAATCGATGCAACGCGAATAAATCTTAACCTT 986
      961 GACATGTAGCGAATTTTCTAGAGTAGATTAAGTCT--TCGGAGACGCTTAACACAGGTG 1017
      987 GACATGTCTGGATTCGGAAGATTTGGAGTGTCTCGCAAGAACGCGAACAAGGTG 1046
      1018 CTGATAGGCTGTCTCAAGTCTGTCTGTGATGATGTTGGGTAAATCCCGCAAGACGCA 1077
      1047 CTGATAGGCTGTCTCAAGTCTGTCTGTGATGATGTTGGGTAAATCCCGCAAGACGCA 1106
      1078 ACCCTTGTCATTAATGGCATCATTTGGGTGGGCACTTTAATGAGACTGCGGTGACAA 1137
      1107 ACCCTTGTCATTAATGGTCTACGAAA-----GGGCACTTAATGAGACTGCGGTGACAA 1161
      1138 CCGAGGAAAGGTGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGCTTCAACG 1197
      1162 CCGAGGAAAGGTGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGCTTCAACG 1221
      1198 TAAATCAATGGCGGTACAGAGGGTGGCCAAACCGCGAGGGGAGGCTAATCTCAGAAAGC 1257
      1222 TCAATCAATGTGTGGGACAGAGGGTGGCCAAACCGCGAGGGGAGGACCAATCCAGAAAC 1281
      1258 GCGTGTAGTCCGGAATCGGAGTGTGCAACTGCACTCCGTGAAGTGGAAATGCTAGTAA 1317
      1282 CGATGTAGTCCGGAATCGGAGTGTGCAACTGCACTCCGTGAAGTGGAAATGCTAGTAA 1341
      1318 CGCGGATCAGCATGTCCGGGTGAATACGTTCCCGGGCTTTGTACACACCGCCCGTCAAC 1377
      1342 CGCGGATCAGCATGTCCGGGTGAATACGTTCCCGGGCTTTGTACACACCGCCCGTCAAC 1401
      1378 CATGGAGTGGGTTTACACCAAGAGAGTGTAAACCTTAAGAGAGGGCGCTTGGCAAG 1437
      1402 CATGGAGTGGGTTTACACCAAGAGTGTAAACCTTAAGAGAGGGCGGATTAACACAG 1461
      1438 TGAGATTCAATGCTGGGGTG 1457
      1462 TAGGATTCAATGCTGGGGTG 1481
```

RESULT 23

```
AA01866
ID   AA01866 standard; DNA; 1526 BP.
XX
AC   AA01866;
XX
DT   16-OCT-2003 (revised)
DT   03-AUG-1999 (first entry)
XX
DE   P.cepacia 16S rRNA gene sequence.
XX
XX   16S rRNA; KK01; primer; PCR; amplification; probe; hybridisation;
XX   detection; diagnosis; ds.
XX
OS   Burkholderia cepacia.
XX
JF07255486-A.
XX
PD   09-OCT-1995.
XX
XX   23-MAR-1994; 94UP-00051739.
XX
XX   23-MAR-1994; 94UP-00051739.
XX
XX   23-MAR-1994; 94UP-00051739.
XX
PA   (CANO ) CANON KK.
XX
DR   WPI; 1995-378541/49.
XX
PT   Pseudomonas cepacia KK01 strain 16S rRNA gene - also related probes and
XX   primers, useful for specific detection of P.cepacia strain KK01.
XX
PS   Claim 1; Page 21; 21pp; Japanese.
XX
XX   This sequence represents the 16S rRNA gene of Pseudomonas cepacia strain
XX   CC KK01. Fragments of the nucleic acid sequence (see AA01872-T02316) are
XX   useful as primers and probes for the specific detection of P.cepacia
XX   strain KK01. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ   Sequence 1526 BP; 382 A; 352 C; 486 G; 306 T; 0 U; 0 Other:
Query Match      78.8%; Score 1148.8; DB 2; Length 1526;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 172; Indels 8; Gaps 2;
QY      1 ATTGAACGCTGGCGGATGCTTTTACATATGCAATGCAAGTCAAGCGCACGCGATGCTTCAT 60
      21 ATTGAACGCTGGCGGATGCTTTTACATATGCAATGCAAGTCAAGCGCACGCGATGCTTCAT 80
      61 CTGTTGGCGAGTGGCGGACGGGTGAGTATGCAATCGGAAGTATCCAGAGAGGGGGGTA 120
      81 CTGTTGGCGAGTGGCGGACGGGTGAGTATGCAATCGGAAGTATCCAGAGAGGGGGGTA 140
      121 ACGCATGAAAGATGTGCTAATACCGCATTAATCTAAGAGAGAAAGCAGGGGATCGAAA 180
      141 GCCCGGGGAAAGCCGATTAATACCGCATTAATCTAAGAGAGAAAGCAGGGGATCGAAA 200
      161 GACCTTCGCTTTTGAAGCGGCGCATGTCTGATTAATGCTAGTTGGTGGGCTTAAAGCCTTAC 240
      201 GGCCTCGCTAATAGGTTGGCCGATGCTGATTAAGTGTGAGGTAAAGGCTTAC 260
      181 GACCTTCGCTTTTGAAGCGGCGCATGTCTGATTAATGCTAGTTGGTGGGCTTAAAGCCTTAC 240
      201 GGCCTCGCTAATAGGTTGGCCGATGCTGATTAAGTGTGAGGTAAAGGCTTAC 260
      241 CAAAGCGACATCATGATGTTGTGTGAGAGAGACACAGCCACACTGGAGCTGAGACAG 300
      261 CAAAGCGACATCATGATGTTGTGTGAGAGAGACACAGCCACACTGGAGCTGAGACAG 320
      301 GCCCAGACTCTTAAGGAGGACGAGTGGGAAATTTTGAACATGGGCGCAAGCCTGATC 360
      321 GCCCAGACTCTTAAGGAGGAGGAGGAGGAGGAAATTTTGAACATGGGCGCAAGCCTGATC 380
      361 CAGCAATGCGCGTGTGAGTGAAGAAAGCCTTGGGTTTGAAGCTTTTCACTCGAAGAA 420
      381 CAGCAATGCGCGTGTGAGTGAAGAAAGCCTTGGGTTTGAAGCACTTTTGTCCGGAAGA 440
      421 AAAGTTACGTTAATATATGCTGACCATGACGCTATGACAGAAAGAACCGGGCTAAC 480
```

```

Db 441 AATCTTGGCTTATATACAGTCGGGGGATGACGGTACCGGAAATTAAGCACCGGCTAAC 500
Qy 481 TAGCTGCACAGCCGGGTAATATACGTAAGGTGCAACCGTAAATCCGAATTAACCTGGGCGT 540
Db 501 TAGCTGCACAGCCGGGTAATATACGTAAGGTGCAACCGTAAATCCGAATTAACCTGGGCGT 560
Qy 541 AAAGGTGCGCAGCGGCTTGTATAGTCAGATGGAATCCCGGGCTTAACTGGGAAT 600
Db 561 AAAGCGTCCGAGCGGTTTGTATAGTCAGATGGAATCCCGGGCTTAACTGGGAAT 620
Qy 601 TGGCTTGAATCTACAAAGCTAGAGTGTGCAGAGGAGGTGGAATTCATGTGTACAG 660
Db 621 TGCATTGTGTACTGGCAAGCTAGAGTATGCGACAGGGGGTAGAATTCACAGTGTACAG 680
Qy 661 TGAATTCGCTAGATATGGAAGAACATCGATGCGGAAGGACGCTCTGGGTTAACT 720
Db 681 TGAATTCGCTAGATATGGAAGAACATCGATGCGGAAGGACGCTCTGGGTTAACT 740
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAATACCTGTAGTCCAGCC 780
Db 741 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAATACCTGTAGTCCAGCC 800
Qy 781 CTAAACGATGTCAACTAGTTGTGGGCTTATTAAGCTTGTGATACGAACTAAACGCTGA 840
Db 801 CTAAACGATGTCAACTAGTTGTGGGCTTATTAAGCTTGTGATACGAACTAAACGCTGA 860
Qy 841 AGTTGACCGCTGGGGAGTACGGTCGCAAGATTAACTCAAGGAATTAACCGGGACCC 900
Db 861 AGTTGACCGCTGGGGAGTACGGTCGCAAGATTAACTCAAGGAATTAACCGGGACCC 920
Qy 901 GCACAAGCGGTGATTTATGTGATTAATTAATGATGCAACGGAAAACTTAACCTTACCCTT 960
Db 921 GCACAAGCGGTGATTTATGTGATTAATTAATGATGCAACGGAAAACTTAACCTTACCCTT 980
Qy 961 GACATGTAGCGAAATTTTCTAGAGATAGATTAGTCT--TCGGAAACGCTAACACAGGTG 1017
Db 981 GACATGTAGCGAAATTTTCTAGAGATAGATTAGTCTGAAAGAAAGAACCGCGCACAGGTG 1040
Qy 1018 CTGCAATGAGCTGTGTGATCGTGTGCTGATGATGTTGGGTTAATGCCGAACAGACCGCA 1077
Db 1041 CTGCAATGAGCTGTGTGATCGTGTGCTGATGATGTTGGGTTAATGCCGAACAGACCGCA 1100
Qy 1078 ACCCTTGTCAATTAATGCTCATCAATTTGGTGGCACTTTAATGAGACTGCGGTGACAAA 1137
Db 1101 ACCCTTGTCTTATGTTGCTAC-----GCAAGAGACACTTAAGAGACTGCGGTGACAAA 1155
Qy 1138 CCGGAGGAGGTGGGATGACGTCAGTCTCATGCGCTTATGGGTAGGGCTTCAACG 1197
Db 1156 CCGGAGGAGGTGGGATGACGTCAGTCTCATGCGCTTATGGGTAGGGCTTCAACG 1215
Qy 1198 TAATTAATGAGCGGCTACAGAGGTTGCCAACCCCGGAGGGGAGCTTAATCTCAAGAAAGC 1257
Db 1216 TCATTAATGAGCGGCTACAGAGGTTGCCAACCCCGGAGGGGAGCTTAATCTCAAGAAAGC 1275
Qy 1258 GCGTGTAGTCCGGATTCGAGTCTGCAACTGCACTCGTGAATTCGGAATCGCTAGTAAT 1317
Db 1276 GCGTGTAGTCCGGATTCGAGTCTGCAACTGCACTCGTGAATTCGGAATCGCTAGTAAT 1335
Qy 1318 CCGGATTCAGCATGTGCGCGGTGAATAGTTCCCGGGCTTTGTACACACCGCCCGGTACAC 1377
Db 1336 CCGGATTCAGCATGTGCGCGGTGAATAGTTCCCGGGCTTTGTACACACCGCCCGGTACAC 1395
Qy 1378 CATGGAGTGGGTTTCAACCAAGACAGTATGTTAACTTAAGAGAGGGCGCTTGCACCG 1437
Db 1396 CATGGAGTGGGTTTCAACCAAGAGTGTAGTCTTAACCGCAAGAGAGAGCGGTACACCG 1455
Qy 1438 TGAGATTCACTGACTGGGGTG 1457
Db 1456 TAGGATTCACTGACTGGGGTG 1475

```

RESULT 24  
AAC86029

```

ID AAC86029 standard; cDNA; 1426 BP.
XX
AC AAC86029;
XX
DT 06-AUG-2003 (revised)
XX
DT 29-AUG-2001 (first entry)
XX
DE 16S rDNA, GCPI12.
XX
KM 16S rDNA; polyphosphate accumulating organism; PAO; probe; primer;
KW detection; phosphorus; waste water; sludge; ss.
XX
OS Unidentified.
XX
PN MO200146459-A1.
XX
PD 28-JUN-2001.
XX
PF 28-DEC-2000; 2000MO-AU001611.
XX
PR 23-DEC-1999; 99AU-00004867.
XX
PA (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.
XX
PI Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ,
XX
DR MPI; 2001-408656/43.
XX
PT Novel oligonucleotide probe or primer useful for detecting polyphosphate
PT accumulating organism in a sample, comprises a sequence that is unique to
PT 16S rDNA of polyphosphate accumulating organisms.
XX
PS Claim 4; Fig 3; 5app; English.
XX
CC The sequences given in AAC86021-30 represent 16S rDNA sequences from
CC polyphosphate accumulating organisms (PAOs). Sequences which are unique
CC to these 16S rDNA sequences are used to create a probe or primer for
CC detecting the relevant organisms. The primer/probe sequences are useful
CC for detecting PAO cells in a sample, by treating cells in the sample to
CC fix cellular contents, contacting fixed cells with the primer/probe which
CC is labelled with a radiolabel, a reporter group or a hapten, under
CC conditions which allow the probe to hybridize with 16S rRNA within the
CC fixed cell, removing unhybridized probe from the fixed cells, and
CC detecting the labeled probe-RNA hybrid by fluorescence in situ
CC hybridization. The primer/probe sequences are useful for identifying PAOs
CC that are capable of biologically removing phosphorus from waste water.
CC Rapid assessment of the presence of a number of PAOs in a waste water
CC sample, can be done using the primer/ probe sequences. They allow quick
CC and convenient assessment of whether a sludge or waste water sample
CC includes PAOs and allows quantitation of PAO cells in samples. (Updated
CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 1426 BP; 361 A; 335 G; 455 C; 275 T; 0 U; 0 Other;
XX
Query Match 78.4%; Score 1142; DB 4; Length 1426;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 1266; Conservative 0; Mismatches 145; Indels 7; Gaps 3;
Qy 12 GCGGATGCTTTACATGCAAGTGAACCGGACGACGAGATCTTGATCTGTGGCGAG 71
Db 11 GCGGATGCTTTACATGCAAGTGAACCGGACGACGAGGAGCAAC--CTGTGTGGCGAG 68
Qy 72 TGGCGGAGCGGTGATTAATGATCGGAACGTAATCCAGAAAGAGGGGGTAAACGATCGAAA 131
Db 69 TGGCGGAGCGGTGATTAAGATTCGGAACGTCCTGGAATGGGGGATTAAGTAACTCGAAA 128
Qy 132 GATGTGCTAATACCGCATATCTTAAGAGGAAAGCAGGGAGATCGAAAGACTTGGCGCT 191
Db 129 GTTACGCTAATACCGCATATCTTGTAGACAGGAAAGCAGGGAGATCGAAAGACTTGGCGTT 188
Qy 192 TTTGAGCGCGCGATGTGATTAAGTGAATGTTGTGGGTAAAGGCTTACCAAGCGACGA 251
Db 189 CGAGGAGCGCGGATGTGATTAAGTGAATGTTGTGGGTAAAGGCTTACCAAGCGACGA 248

```

```
QY 252 TCAAGTAGTGGTCTGAGAGACGACGACCACTGAGACTGAGACACGCGCCAGCTCC 311
DB 249 TCCGTAGCGGGTCTGAGAGATGATCCGCCACACTGGAATGAGACACGCGTCCAGATCC 308
QY 312 TACGGAGGCGACGAGTGGGGAATTTTGGACATGGGGCGCAAGCCTGATCCAGGAATGCG 371
DB 309 TACGGAGGCGACGAGTGGGGAATTTTGGACATGGGGCGCAAGCCTGATCCAGCATCCG 368
QY 372 CGTAGAGGAAGAGGCGCTTGGGTTGTAAAGCTCTTTCAGTCGAGAAAGAAAGGTTACGG 431
DB 369 CGTAGAGGAAGAGGCGCTTGGGTTGTAAAGCTCTTTCGCGGGAGAAATCCACAGGG 428
QY 432 TAAATTAATCGTAGCCATGACGATATCGACAGAAAGACACCGGCTTAATGAGTGCAGC 491
DB 429 TAAATTAATCGTAGATGACGATACCGGAATAGAAAGACCGGCTTAATGAGTGCAGC 488
QY 492 AGCCGCGGTAAATCGTAGGAGTGCAGAGGCTTAATGGAATTAATGAGGCTTAAGGTCGC 551
DB 489 AGCCGCGGTAAATCGTAGGAGTGCAGAGGCTTAATGGAATTAATGAGGCTTAAGGTCGC 548
QY 552 AGGCGGCTTGTAGTCAAGTGAATCCCGGGCTTAACCTGGGAATTGCGTTTGA 611
DB 549 AGGCGGCTTGTAGTCAAGTGAATCCCGGGCTTAACCTGGGAATTGCGTTTGA 608
QY 612 CTACAAAGCTAGAGTGTGGCAGAGGAGTGAATTCATGATGTAGCAGTGAATGCGTA 671
DB 609 CTGCAAGCTGTGAGTGTGGCAGAGGAGTGAATTCATGATGTAGCAGTGAATGCGTA 668
QY 672 GAGATATGGAAGAACATGATGCGAGAGGACGCTCTGGGTTAACTGACGCTCATGC 731
DB 669 GAGATATGGAAGAACACCGATGCGAGAGGACGCTCTGGGTTAACTGACGCTCATGC 728
QY 732 ACGAAAGCGTGGGAGCAACAGAGTTAGATACCTGGTAGTCCACGCGCTTAAGCATGT 791
DB 729 ACGAAAGCGTGGGAGCAACAGAGTTAGATACCTGGTAGTCCACGCGCTTAAGCATGT 788
QY 792 CAACCTAGTGTGTGG--GCCTTATTAGGCTTGTGAACGAGCTTAACGCGTGAAGTTGACG 849
DB 789 CAACCTAGTGTGTGGAGGAGTTAAACCTTTTGTAGTCCGTAACGCGTGAAGTTGACG 848
QY 850 CTTGGGAGTACGCTGCGAAGATTAACTCAAGAGAAATTGACGCGGACCCGACACACG 909
DB 849 CTTGGGAGTACGCGCGCGCAAGGCTTAACCTCAAGAGAAATTGACGCGGACCCGACACACG 908
QY 910 GTGATTAATGAGATTAATTCGATGCAACGGAACAACTTAACCTTGAACATGTAG 969
DB 909 GTGATTAATGAGATTAATTCGATGCAACGGAACAACTTAACCTTGAACATGTAG 968
QY 970 CGAATTTTCTAGAGATGATTAGTCT---TCGGGAACGCTTAACAGAGTGTGATGGC 1026
DB 969 GGAATCCCGAGAGATTTTGGAGTCTCGCAAGAGAACCTGAACACAGTGTGATGGC 1028
QY 1027 TGTGCTACGCTGCTGCTGAGATGTTGAGTTCGCGCAACGAGCGCAACCTTGTCTC 1086
DB 1029 TGTGCTACGCTGCTGCTGAGATGTTGAGTTCGCGCAACGAGCGCAACCTTGTCTC 1088
QY 1087 AATTAATGCAATATTTGTTGGGCACTTTAATGACATGCGCGGAGCAAAACCGGAGAA 1146
DB 1089 AATTAATGCAATATTTGTTGGGCACTTTAATGACATGCGCGGAGCAAAACCGGAGAA 1148
QY 1147 GGTGGGAGTGAAGTCAAGTCTCATGCGCTTAATGGGTAGGCGCTTCAACGTAATACAT 1206
DB 1149 GGTGGGAGTGAAGTCAAGTCTCATGCGCTTAATGGGTAGGCGCTTCAACGTAATACAT 1208
QY 1207 GCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTTAATTCAGAAAGCGCGTGTAG 1266
DB 1209 GGTGCGTCCAGAGGTTGCAACCCGCGAGGGGAGCAATCCAGAAAGCGCATGTAG 1268
QY 1267 TCCGGAATGGAATCTGCAACTGATCGTGAAGTCCGAAATGCGATGAATGCGGAGTCA 1326
DB 1269 TCCGGAATGGAATCTGCAACTGATCGTGAAGTCCGAAATGCGATGAATGCGGAGTCA 1328
```

```
QY 1327 GCATGTCGCGGTGAATACGTTCCCGGATCTTGTACACACCGCCCTTACACATGGGAGT 1386
DB 1329 GCATGTCGCGGTGAATACGTTCCCGGATCTTGTACACACCGCCCTTACACATGGGAGT 1388
QY 1387 GGGTTTACCCAGAGAGGATGTTAACTGTAAGAGG 1424
DB 1389 GGGTTTACCCAGAGAGTATGTTAGCTTAACCGCAAGAGG 1426

RESULT 25
ADQ16355
ID ADQ16355 standard; DNA; 1496 BP.
AC ADQ16355;
DT 09-SEP-2004 (first entry)
DE Nucleotide sequence of a Bordetella variable 16S rRNA gene region.
KW fragmentation-based method; mass spectrometric method;
KW nucleic acid polymorphism; nucleic acid mutation; genetic disease;
KW chromosome abnormality; obesity; atherosclerosis; cancer; haplotype; ss;
KW rRNA gene.
OS Bordetella petrii.
PN W02004050839-A2.
PD 17-JUN-2004.
XX 26-NOV-2003; 2003MO-US037931.
XX 27-NOV-2002; 2002US-0429895P.
XX (SEQU-) SEQUENOM INC.
PI Van Den Boom D, Boecker S;
XX WPI; 2004-487567/46.
DR Use of fragmentation-based methods and systems, e.g. mass spectrometric
PT methods for the analysis of sequence variations including nucleic acid
PT polymorphisms and mutations.
XX Example 4; SEQ ID NO 32; 198bp; English.
XX
CC The specification describes the use of fragmentation-based methods and
CC systems including mass spectrometric methods for the analysis of sequence
CC variations including nucleic acid polymorphisms and mutations. The
CC fragmentation-based methods and systems of the invention are useful for
CC the analysis of sequence variations including nucleic acid polymorphisms
CC and mutations. The methods are useful for identifying a genetic disease
CC or chromosome abnormality; identifying a predisposition to a disease or
CC condition including obesity, atherosclerosis, or cancer; identifying an
CC infection by an infectious agent; providing information relating to
CC identity, heredity, or histocompatibility; identifying pathogens; or
CC determining haplotypes. ADQ16353-ADQ16361 represent Bordetella variable
CC 16S rRNA gene regions. Amplicons from this region were used to demonstrate
CC the invention. They were used to demonstrate a method for bacterial
CC typing by base-specific fragmentation.
XX
SQ Sequence 1496 BP; 372 A; 355 C; 475 G; 293 T; 0 U; 1 Other;
Query Match 78.3%; Score 1140.6; DB 12; Length 1496;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1279; Conservative 0; Mismatches 165; Indels 10; Gaps 3;
```

```
QY 7 CGTGGCGGATGCTTTTACATCAAGTCAACGCGACGCGGATGCTTGCATCTGCTG 66
DB 1 CGTGGCGGATGCTTTTACATCAAGTCAACGCGACGCGGATGCTTGCATCTGCTG 58
QY 67 GCGATGCGCGGACGCGGTGAGTAATGATCGGAAGTATCCAGAGAGGCGGATACGAT 126
```

Db 59 GCGAGTGGCGAAGCGGATGAGTAATGTATCGGAAAGTGCACGATAGCGGGGATTAATACGTACG 118  
Qy 127 CGAAAGATGTCTAATATCCGCATTAATCTTAAGAGAGAAACAGGGGATGAAAGACCTT 186  
Db 119 CGAAAGCTTAGTAATATCCGCATTAATCGGAGGAGAAAGCGGGGATCTTGGGGCTTC 178  
Qy 187 GGCCTTTTGGAGGCGGCATGTCTGATTAGCTAGTTGGTGGGTTAAAGGCTTACCAAGGC 246  
Db 179 GCACTATTGGAGGCGGCATGATGAGTTAGCTAGTTGGTGGGTTAAAGGCTTACCAAGGC 238  
Qy 247 GACGATCAGTAGTTGGTCTGAGAGAGACGACCACTGGGATCTGAGACACGGCCAG 306  
Db 239 GAGGATCCGTAGTGGTTTGAAGAGACGACCACTGGGATCTGAGACACGGCCAG 298  
Qy 307 ACTTCCTACGGGAGGCGAGCATGTGGGAAATTTTGGACAAATGGGCGCAAGCTTATCCAGCA 366  
Db 299 ACTTCCTACGGGAGGCGAGCATGTGGGAAATTTTGGACAAATGGGCGCAAGCTTATCCAGCA 358  
Qy 367 TGGCGGCTGATGAGAAAGGCTTGGGTTGTAAAGCTTTTCAAGTGCAGAGAAAGGT 426  
Db 359 TCCGCGCTGTGCGATGAAAGGCTTGGGTTGTAAAGCACTTTTGGCAGAGAAAGAGCGC 418  
Qy 427 TACGTAATTAATCTGTATGATGACCGATGACGATATCGACAGAAAGACACCGGCTTAATCTAGTG 486  
Db 419 TCTGGCTAATATACCTGGGCAACTGACGCTACCTGACAGAAATTAAGCACCGGCTTAATCTAGTG 478  
Qy 487 CCAGCAGCCCGGCTTAATCTAGAGGTGCAAGCGTAACTCGGAATTAATCTGGGCGCTTAAAGGG 546  
Db 479 CCAGCAGCCCGGCTTAATCTAGAGGTGCAAGCGTAACTCGGAATTAATCTGGGCGCTTAAAGGG 538  
Qy 547 TGGCAGGCGGCTTGTAAATGATGTAATTCGCCGCGCTTAACTCGGAAATTCGCTT 606  
Db 539 TGGCAGGCGGCTTGTAAATGATGTAATTCGCCGCGCTTAACTCGGAAATTCGCTT 598  
Qy 607 TGAACCTACAAAGCTTAAGTGTGCGAGAGGAGGTGAATTCATGTGTAGACAGTGAAT 666  
Db 599 TTTAATCTACCGGCTTAAGTGTGCGAGAGGAGGTGAATTCGCCGCTTGTAGAGTGAAT 658  
Qy 667 GCGTAAGATTAATGAAAGAAATCATGTGCGAAGCGCCCTCGGGTTAACTATGACGT 726  
Db 659 GCGTAAGATTAATGAAAGAAATCATGTGCGAAGCGCCCTCGGGTTAACTATGACGT 718  
Qy 727 CATGACGAAAGGCTGGGAGAGCAAAACAGATTAATCTCGGTAGTCCAGCCCTTAAC 786  
Db 719 CATGACGAAAGGCTGGGAGAGCAAAACAGATTAATCTCGGTAGTCCAGCCCTTAAC 778  
Qy 787 GATGTCAACTAGTTGTTGGGCTTAATTAAGCTTGTAAAGAGCTTAACCGGTGAAGTTGA 846  
Db 779 GATGTCAACTAGTTGTTGGGCTTAATTAAGCTTGTAAAGAGCTTAACCGGTGAAGTTGA 838  
Qy 847 CCGCCTGGGAGATACGCTCGCAAGATTAATCTAAAGAAATTAAGCGGGGACCCGACAA 906  
Db 839 CCGCCTGGGAGATACGCTCGCAAGATTAATCTAAAGAAATTAAGCGGGGACCCGACAA 898  
Qy 907 GCGGTGATTAATGATTAATCGATGCAAGCGAAACCTTAACCTACCTTTGACATG 966  
Db 899 GCGGTGATTAATGATTAATCGATGCAAGCGAAACCTTAACCTACCTTTGACATG 958  
Qy 967 TACGGAATTTTCTAGAGATAGATTAAGTCT--TCGGGAAACGCTTAACACAGGTCTGCAT 1023  
Db 959 TCTGGAATGCGAAAGAGATTTGGCAGTGCTCGCAAGAGAACCGGAACACAGGTCTGCAT 1018  
Qy 1024 GGTGTCTGCAAGCTGCTGTGTGAGATTTGGGTTAAGTCCCGCAACGAGCGCAACCTT 1083  
Db 1019 GGTGTCTGCAAGCTGCTGTGTGAGATTTGGGTTAAGTCCCGCAACGAGCGCAACCTT 1078  
Qy 1084 GTCAATTAATGATCACTTTGTTGGGCACTTAATAGAGCTCCGGTGAACAAACGGAG 1143  
Db 1079 GTCAATTAATGATCACTTTGTTGGGCACTTTAGAGCTCCGGTGAACAAACGGAG 1133  
Qy 1144 GAAGGTGGGAGATGACGTCAAGTCTCAATGGCCTTATGGGTAGGGCTTCAACGTAATAC 1203  
Db 1134 GAAGGTGGGAGATGACGTCAAGTCTCAATGGCCTTATGGGTAGGGCTTCAACGTAATAC 1193

Qy 1204 AATGGCGGTACAGAGGCTTGGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGGCTGC 1263  
Db 1194 AATGGTCGAGACAGAGGCTTGGCAACCCGCAAGGGGAGCCAAATCCAGAAACCGGATGC 1253  
Qy 1264 TAGTCGGATCGGAGTCTGCAACTGCACTCCGTGAATGCGAATGCTTAATGCGGGA 1323  
Db 1254 TAGTCGGATCGGAGTCTGCAACTGCACTCCGTGAATGCGAATGCTTAATGCGGGA 1313  
Qy 1324 TCAGCATGTGCGGCTGAATACGTTCCGGGCTTTGTACACACCCGCCGTCAACCATGGG 1383  
Db 1314 TCAGCATGTGCGGCTGAATACGTTCCGGGCTTTGTACACACCCGCCGTCAACCATGGG 1373  
Qy 1384 AGTGGGTTTACACAGAGGATGATCTAACCGTAAGAGAGGGGCTTGCACAGTGAAT 1443  
Db 1374 AGTGGGTTTACACAGAGGATGATCTAACCGTAAGAGAGGGGCTTGCACAGTGAAT 1433  
Qy 1444 TCATGACTGGGGTG 1457  
Db 1434 TCATGACTGGGGTG 1447

RESULT 26  
AEA01071  
ID AEA01071 standard; DNA; 1496 BP.  
XX  
AC AEA01071:  
XX  
XX 28-JUL-2005 (first entry)  
XX  
XX  
XX Bordetella petril 16S rRNA gene SEQ ID NO:32.  
DE  
XX mass spectroscopy; analysis; polymorphism; base-specific fragmentation;  
KM 16S rRNA; 16S ribosomal RNA; ds.  
XX  
XX Bordetella petril.  
OS  
XX  
XX US2005112590-A1.  
PN  
XX  
XX 26-MAY-2005.  
PD  
XX  
XX 26-NOV-2003; 2003US-00723365.  
PF  
XX  
XX 27-NOV-2002; 2002US-0429895P.  
PR 25-APR-2003; 2003US-0466006P.  
PR  
XX  
XX (VBOC/) VAN DEN BOOM D.  
PA (BOEC/) BOECKER S.  
XX  
XX Van Den Boom D, Boecker S;  
PI  
XX  
XX WPI, 2005-371655/38.  
DR  
XX  
XX Determining sequence variations in a target biomolecule or nucleic acid  
PT molecule, useful for analyzing sequence variations or polymorphisms in  
PT hemophilia, Alzheimer's disease, Cystic Fibrosis, and Down's syndrome and  
PT diabetes.  
XX  
XX Example 4; SEQ ID NO 32; 78bp; English.  
PS  
XX  
XX The invention relates to a method for determining sequence variations in  
CC a target biomolecule or nucleic acid molecule. The method comprises  
CC cleaving the target biomolecule into fragments by contacting the target  
CC biomolecule or nucleic acid molecule with one or more specific cleavage  
CC reagents, cleaving or simulating cleavage of a reference biomolecule or  
CC nucleic acid molecule into fragments with the same cleavage reagents,  
CC determining mass signals of the fragments produced, determining  
CC differences in the mass signals between the fragments produced, and  
CC determining a reduced set of sequence variation candidates from the  
CC differences in the mass signals and determining sequence variations in  
CC the target compared to the reference biomolecule or nucleic acid  
CC molecule. The methods and compositions of the present invention are  
CC useful for analyzing sequence variations and/or polymorphisms in diseases

CC such as hemophilia, thalassemia, Duchenne Muscular Dystrophy,  
CC Huntington's Disease, Alzheimer's Disease and Cystic Fibrosis, and  
CC Trisomy 21 (Down's Syndrome), Trisomy 13 (Patau Syndrome), Trisomy 18  
CC (Edward's Syndrome), Monosomy X (Turner's Syndrome) and other sex  
CC chromosome aneuploidies such as Klinefelter's Syndrome (XXY), and  
CC diabetes, arteriosclerosis, obesity, various autoimmune diseases and  
CC cancer (e.g., colorectal, breast, ovarian, lung). The present sequence  
CC represents a 16S rRNA gene sequence, which is used in an example from the  
CC present invention for the bacterial typing by base-specific  
CC fragmentation.

XX Sequence 1496 BP; 372 A; 355 C; 475 G; 293 T; 0 U; 1 Other;

Query Match 78.3%; Score 1140.6; DB 14; Length 1496;

Best Local Similarity 88.0%; Pred. No. 0;

Matches 1279; Conservative 0; Mismatches 165; Indels 10; Gaps 3;

```
QY 7 CGCTGGCGGCGGCTTTTACATCATGCAAGTCGACGCGACGCGATGCTTGATCTGTG 66
DB 1 CGTAGCGGGGATGCTTTACATCATGCAAGTCGACGCGACGCGA--CTTGGCTTGGCG 58
QY 67 GCGAGTGGCGGACGGGTGATGATCGAATCGGACGTAATCCAGAGAGGGGGTAAACGAT 126
DB 59 GCGAGTGGCGGACGGGTGATGATCGAATCGGACGTAATCCAGAGAGGGGGTAAACGAT 118
QY 127 CGAAGATGTCCTAATACCGCATATCTTAAGAGAGAAACAGGGGATCGAAGACCTT 186
DB 119 CGAAGATGTCCTAATACCGCATATCGCTTAAGAGAGAAACAGGGGATCGAAGACCTT 178
QY 187 GCGCTTTTGAAGCGGCGCATGCTGATTTAGTGGTGGGTTAAAGGCTTACCAAGGC 246
DB 179 GCATATTGAGCGGCGCATGCTGATTTAGTGGTGGGTTAAAGGCTTACCAAGGC 238
QY 247 GACCATGATGTTGTTGTTGAGAGAGACGACGACGACGCTGGAGACGCGCCAG 306
DB 239 GACCATGATGTTGTTGTTGAGAGAGACGACGACGACGCTGGAGACGCGCCAG 298
QY 307 ACTCTTACGGAGGACGATGAGGAGATTTTGAACAATGGGCGCAAGCTGATCCAGCA 366
DB 299 ACTCTTACGGAGGACGATGAGGAGATTTTGAACAATGGGCGCAAGCTGATCCAGCA 358
QY 367 TGCCGCGTGAAGAGAGGCTTTCGGGTTGTAAGCTCTTCACTGCAAGAAAGT 426
DB 359 TGCCGCGTGAAGAGAGGCTTTCGGGTTGTAAGCTCTTCACTGCAAGAAAGT 418
QY 427 TACGTAATTAATCTGTAACCATGACGATTCGACAGAGAGACGCGCTAACTACG 486
DB 419 TCTGCTAATTAATCTGTAACCATGACGATTCGACAGAGAGACGCGCTAACTACG 478
QY 487 CCAGCAGCGCGGTAAATAGTAGGTTGCAAGGCTTAATCGAATTAATCGGCGTAAAGG 546
DB 479 CCAGCAGCGCGGTAAATAGTAGGTTGCAAGGCTTAATCGAATTAATCGGCGTAAAGG 538
QY 547 TGCGGAGCGGCTTTGTAAGTCAATGTGAATTCCTCGGCTTAACTTGGGATTCG 606
DB 539 TGCGGAGCGGCTTTGTAAGTCAATGTGAATTCCTCGGCTTAACTTGGGATTCG 598
QY 607 TGAATCTCAAAAGCTAGAGTGGCGAGGAGGAGGAAATTCAGATGTAAGAGAGAAAT 666
DB 599 TTTAATCTAATCGGCTTGAAGTGTGTCAGAGGAGGAGGAAATTCAGATGTAAGAG 658
QY 667 GCGTAGAGATTAAGAAACATCGATGCGGAGAGGACGCTCTTGGGTTAACTACGCT 726
DB 659 GCGTAGAGATTAAGAAACATCGATGCGGAGAGGACGCTCTTGGGTTAACTACGCT 718
QY 727 CATGACGAAAGCGTGGGAGCAACAGATTAGATACCTTGGTATGTCACGCTTAAAC 786
DB 719 CATGACGAAAGCGTGGGAGCAACAGATTAGATACCTTGGTATGTCACGCTTAAAC 778
QY 787 GATGTCACCTAGTGTGGGCTTATTAAGCTTGTGAACGAGCTTAAGCGGTAAGTTGA 846
DB 779 GATGTCACCTAGTGTGGGCTTATTAAGCTTGTGTGAGGACGAGCTTAAGCGGTAAGTTGA 838
```

```
QY 847 CGGCTGGGGAGTACGGTCCGAAGATTAAACTCAAGAAATTGACGGGACCCGACAA 906
DB 839 CGGCTGGGGAGTACGGTCCGAAGATTAAACTCAAGAAATTGACGGGACCCGACAA 898
QY 907 GCGGTGGATTAATGATTAATTCATGCAACGCGAAGAACTTACCTACCTTTGACATG 966
DB 899 GCGGTGGATTAATGATTAATTCATGCAACGCGAAGAACTTACCTACCTTTGACATG 958
QY 967 TACGGAATTTTCAAGATTAATGATGCT--TGGGAACGCTAACACAGGTCGTCAT 1023
DB 959 TCTGAATGCGGAAGATTGGCAGTCTCGCAAGAGAACCGGAACACAGTGTGCAT 1018
QY 1024 GCGTCTGTCAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAAGACGCAACCTT 1083
DB 1019 GCGTCTGTCAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAAGACGCAACCTT 1078
QY 1084 GTCAATTAATGCAATCATTTTGGTGGGACCTTAATGAGACTCGCGTGAACAAACGGAG 1143
DB 1079 GTCAATTAATGCAATCATTTTGGTGGGACCTTAATGAGACTCGCGTGAACAAACGGAG 1133
QY 1144 GAAGTGGGATGACGTCAGATCTCATGACCGCTTATGGGTAGGCTTCAACGTAATAC 1203
DB 1134 GAAGTGGGATGACGTCAGATCTCATGACCGCTTATGGGTAGGCTTCAACGTCATAC 1193
QY 1204 AATGGCGCGTACAGAGGGTTGCCAACCGCGAGGGGAGCTTAATCTCAGAAAGCGCTCG 1263
DB 1194 AATGGTGGGACAGAGGGCTGCGAACCGCGAAGGGGAGCCCATCCAGAAACCCGATCG 1253
QY 1264 TAGTCCGATTCGAGTCTGCAACTGCACTCGTGAATCGGAATCGTAAATCGCGGA 1323
DB 1254 TAGTCCGATTCGAGTCTGCAACTGCACTCGTGAATCGGAATCGTAAATCGCGGA 1313
QY 1324 TCAGCATGTGCGGTTAATGCTTCCGGGCTTTGTAACAACCGCCGCTCAACCATACGG 1383
DB 1314 TCAGCATGTGCGGTTAATGCTTCCGGGCTTTGTAACAACCGCCGCTCAACCATACGG 1373
QY 1384 AGTGGCTTTCACCAAGACGATGTAACCGTAAGAGAGGCGCTTGCCAGGTAGAT 1443
DB 1374 AGTGGCTTTCACCAAGAGTGTAGTGAACCGTAAGAGAGGCGCATTAACAGGTAGAT 1433
QY 1444 TCATGACTGGGGTG 1457
DB 1434 TCATGACTGGGGTG 1447

RESULT 27
AA064008
ID AA064008 standard; rRNA; 1532 BP.
XX
AC AA064008;
XX
XX 27-AUG-2003 (revised)
DT 22-JUN-1994 (first entry)
XX
DE 16S rRNA gene.
XX
KM 16S rRNA; probe; detection; procine atrophic rhinitis; hybridisation;
KM Bordetella bronchiseptica; pig raising; ss.
XX
OS Bordetella bronchiseptica.
XX
XX JP05336999-A.
XX
XX 21-DEC-1993.
XX
XX 10-JUN-1992; 92JP-00150688.
XX
XX 10-JUN-1992; 92JP-00150688.
XX
XX (NISE-) NIHON SEIFUN KK.
XX (ZENK-) ZENKOKU NOGYO KD RENGOKAI.
XX
XX WPI, 1994-037379/05.
DR
```



XX B. bronchiseptica 16S rRNA fragments - used as probes in the detection of  
PT porcine atrophic rhinitis.  
XX  
XX Claim 1; Page 8-9; 12pp; Japanese.  
XX  
CC DNA sequences (AA064009-064031) are fragments of the 16S rRNA gene from  
CC B. bronchiseptica (AA055187). The fragments are used as probes to detect  
CC porcine atrophic rhinitis caused by the Bordetella bronchiseptica  
CC bacterium. Also claimed are 3 DNA fragments complementary to the 436-466  
CC region of the 16S rRNA (AA064032-034). A specific DNA sequence from the 51  
CC rRNA was selected and 2 probes were designed (AA064035 and AA064039) for  
CC the detection of B. bronchiseptica. Primers (AA064036-37) were used to  
CC clone the 16S gene. Sequences (AA064034) is the preferred probe used in  
CC the detection process. (Updated on 27-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 1532 BP; 376 A; 365 C; 490 G; 301 T; 0 U; 0 Other;  
Query Match 78.1%; Score 1138; DB 2; Length 1532;  
Best Local Similarity 87.7%; Pred. No. 0;  
Matches 1280; Conservative 0; Mismatches 170; Indels 10; Gaps 3;  
QY 1 ATTGAAGCTGCGCGGACGCTTTACACATGCAAGTCGAACGGCAGCAGATGCTTGAT 60  
DB 29 ATTGAAGCTGCGCGGACGCTTTACACATGCAAGTCGAACGGCAGCAGATGCTTGAT 86  
QY 61 CTGCTGCGAGTGGCGGACGGGTGATATGATCGAAGCTATCCAGAAAGAGGGGGTGA 120  
DB 87 CTGCTGCGAGTGGCGGACGGGTGATATGATCGAAGCTATCCAGAAAGAGGGGGTGA 146  
QY 121 AGCGATGAAAGATGTGCTAATACCGCATATACCTTAAGAGGAAAGCAGGGGATCGAA 180  
DB 147 ACTACGGAAAGGTGGCTAATACCGCATACCGCTTACGGGGAAAGCGGGGACCTTCG 206  
QY 181 GACCTTGCGCTTTTGAGCGCGCGATGCTGATTAAGTATGATGATGAGGCTTAC 240  
DB 207 GCGCTGCGCACTATTGAGCGCGCGATGCTGATTAAGTATGATGATGAGGCTTAC 266  
QY 241 CAAGGCGACGATCAATGTTGTTGTTGAGAGAGCAACGACCACTGGACTGAGACAG 300  
DB 267 CAAGGCGACGATCAATGTTGTTGTTGAGAGAGCAACGACCACTGGACTGAGACAG 326  
QY 301 GCCCAGACTCCTACGGGAGGACGAGTGGGAAATTTTGAACAATGGCGGCAAGCTGATC 360  
DB 327 GCCCAGACTCCTACGGGAGGACGAGTGGGAAATTTTGAACAATGGCGGCAAGCTGATC 386  
QY 361 CAGCAATGCGCGGTGATGAAGAGGCTTCGAGTTTAAAGCTTTTCACTGAGAGAA 420  
DB 387 CAGCAATGCGCGGTGATGAAGAGGCTTCGAGTTTAAAGCTTTTCACTGAGAGAA 446  
QY 421 AAAAGTTACGTTAATATGTCGATGACCATGACCGTATCGACAGAAAGACCGGGCTAAC 480  
DB 447 AACGGCAACGGGCTAATATCTGTGCAACTGACCGTATCTCAAGATTAAGACCGGGCTAAC 506  
QY 481 TAGCTGCACAGCGCGGTAAATGATGATGAGGTCGAACTGTAATCGAATTTTACTGGGCGT 540  
DB 507 TAGCTGCACAGCGCGGTAAATGATGATGAGGTCGAACTGTAATCGAATTTTACTGGGCGT 566  
QY 541 AAAAGGTGCGCAGCGCGGCTTGTAAATGATGATGAGGTCGAACTGTAATCGAATTTTACTGGGCGT 600  
DB 567 AAAAGGTGCGCAGCGCGGCTTGTAAATGATGATGAGGTCGAACTGTAATCGAATTTTACTGGGCGT 626  
QY 601 TGGCTTTGAAACTCAAAAGCTAAGTGTGCAAGGAGGAGTGAATTCATGTGTAGACG 660  
DB 627 TGGCTTTGAAACTCAAAAGCTAAGTGTGCAAGGAGGAGTGAATTCATGTGTAGACG 686  
QY 661 TGAATGCGTGAATATGAAAGACATCGATGGCGAAGGACGCTCGTGGTTAACT 720  
DB 687 TGAATGCGTGAATATGAAAGACATCGATGGCGAAGGACGCTCGTGGTTAACT 746  
QY 721 GACGCTCATGACGAAAGCTGAGGAGCAAAAGATTAATACCTGTATGTCACGCC 780  
DB 747 GACGCTCATGACGAAAGCTGAGGAGCAAAAGATTAATACCTGTATGTCACGCC 806

QY 781 CTAACGATGTCAACTAGTGTGGGCTTATTAGGCTTGTAAAGCAAGCTAACGGCTGA 840  
DB 807 CTAACGATGTCAACTAGTGTGGGCTTGTAAAGCAAGCTAACGGCTGA 866  
QY 841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTTAAACTCAAGATTTGACGGGAGCC 900  
DB 867 AGTTGACCGCTGGGAGTACGCTCGCAAGATTTAAACTCAAGATTTGACGGGAGCC 926  
QY 901 GCACAAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAACCTTACCTACCTT 960  
DB 927 GCACAAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAACCTTACCTACCTT 986  
QY 961 GACATGACGAATTTTCTAGAGATAGATTAGTCT---TCGGGAAAGCTTACACAGGTG 1017  
DB 987 GACATGACGAATTTTCTAGAGATAGATTAGTCT---TCGGGAAAGCTTACACAGGTG 1046  
QY 1018 CTGCATGCTGTGCTGACGCTGCTGTGAGATGTTGGGTTAAGTCCGCAACGACGCA 1077  
DB 1047 CTGCATGCTGTGCTGACGCTGCTGTGAGATGTTGGGTTAAGTCCGCAACGACGCA 1106  
QY 1078 ACCCTTGCTAATTAATGCTCAATCTTTGGTTGGGCACTTTAATGAGACTGCGGTGACAA 1137  
DB 1107 ACCCTTGCTAATTAATGCTCAATCTTTGGTTGGGCACTTTAATGAGACTGCGGTGACAA 1161  
QY 1138 CCGGAGAAAGGTGGGAGTACGCTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTCACAG 1197  
DB 1162 CCGGAGAAAGGTGGGAGTACGCTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTCACAG 1221  
QY 1198 TATACAAATGCGCGCTGACAGAGGTTGCCAACCCGAGAGGGGAGCTAATCTCAGAAAGC 1257  
DB 1222 TATACAAATGCGCGCTGACAGAGGTTGCCAACCCGAGAGGGGAGCTAATCTCAGAAAGC 1281  
QY 1258 GCGTGTAGTCCGATCGAGATGCTGCACTCGACTCCGTGAAGTCCGAATTCGTATGAT 1317  
DB 1282 GCGTGTAGTCCGATCGAGATGCTGCACTCGACTCCGTGAAGTCCGAATTCGTATGAT 1341  
QY 1318 CCGGATGACGATGTTGGGAGTGAATGCTTCCGGGCTTTGTACACACCGCCGTCACAC 1377  
DB 1342 CCGGATGACGATGTTGGGAGTGAATGCTTCCGGGCTTTGTACACACCGCCGTCACAC 1401  
QY 1378 CATGGAGTGGGTTTCCACGAGAGCAGTATCTAACCGTAAAGAGGGGCTTGGCCACGG 1437  
DB 1402 CATGGAGTGGGTTTCCACGAGAGTATGTTAGCTTAACCGCAAGGGGGGAGTTACACGG 1461  
QY 1438 TGAATTCATGACTGGGCTG 1457  
DB 1462 TGAATTCATGACTGGGCTG 1481  
RESULT 28  
ADM12666  
ID ADM12666 standard; DNA; 1530 BP.  
XX  
XX ADM12666;  
XX  
DT 07-APR-2005 (first entry)  
XX  
DE Varivoxrax paradoxus 16S rDNA, SEQ ID NO:1.  
XX  
KM Stereoselective synthesis; enantiomeric enrichment;  
KM beta-amino acid synthesis; cell culture; 16S ribosomal RNA; 16S rRNA; de.  
OS Varivoxrax paradoxus.  
XX  
PN US2005009151-A1.  
XX  
PD 13-JAN-2005.  
XX  
PF 22-JUN-2004; 2004US-00875161.  
XX  
PR 10-JUL-2003; 2003US-0486032P.  
PR 02-SEP-2003; 2003US-0499622P.





AAV24295  
ID AAV24295 standard; DNA; 1535 BP.  
XX  
XX AAV24295;  
AC  
XX  
XX 14-SEP-1998 (first entry)  
XX  
XX Burkholderia cepacia 16S ribosomal RNA gene.  
DE  
XX  
XX Antibacterial; antimycobacterial; oligonucleotide; infection; therapy;  
KM ribosome binding site; Shine-Dalgarno; ribosomal RNA; cystic fibrosis;  
KM tuberculosis; ss.  
XX  
OS Burkholderia cepacia.  
XX  
XX WO9814567-A2.  
XX  
XX 09-APR-1998.  
XX  
XX 30-SEP-1997; 97MO-US018094.  
XX  
XX 01-OCT-1996; 96US-0027729P.  
XX  
XX (ADRE-) ADVANCED RES & TECHNOLOGY INST.  
XX  
XX Martin WJ, Winiowski P;  
XX  
XX WPI; 1998-240079/21.  
XX  
XX Use of oligo:nucleotide(s) corresponding to bacterial 16S rRNA - for  
PT inhibiting bacterial protein expression and treating bacterial infection.  
XX  
XX Claim 26; Page 62-63; 73pp; English.  
XX  
XX This polynucleotide comprises the 16S ribosomal RNA (rRNA) gene of  
CC Burkholderia cepacia. The invention relates to methods and compositions  
CC for the treatment of Gram-negative bacterial infections employing novel  
CC oligonucleotides as antimicrobial agents. The oligonucleotides are  
CC targeted to the Shine-Dalgarno (SD) region of prokaryotes to inhibit  
CC bacterial expression and hence inhibit bacterial infection. They  
CC preferably comprise 10-35 consecutive bases of the 3' end of a bacterial  
CC 16S rRNA (see also AAV24291-94). An oligonucleotide may also include a  
CC transport moiety and may have DNA phosphate modifications to increase  
CC nuclease resistance, or may be formulated in a liposome. A claimed method  
CC for treating a bacterial infection of a patient comprises administering a  
CC liposomal formulation of such an oligonucleotide. The oligonucleotides  
CC can be used particularly for treating bacterial infections in pulmonary  
CC diseases such as cystic fibrosis or tuberculosis. Since the SD sequence  
CC is not present in eukaryotic cells, the oligonucleotides provide a  
CC pathogen-specific therapeutic method  
XX  
XX Sequence 1535 BP; 385 A; 355 C; 488 G; 307 T; 0 U; 0 Other;  
SQ  
Query Match 77.9%; Score 1135.2; DB 2; Length 1535;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 1278; Conservative 0; Mismatches 173; Indels 9; Gaps 3;  
QY 1 ATTGAACGCTGGCGGCGCTGTTTACATGCAAGTGAACGCGACGCGATGCTTGAT 60  
DB 29 ATTGAACGCTGGCGGCGCTGTTTACATGCAAGTGAACGCGACGCGATGCTTGAT 88  
QY 61 CTGGTGGCGATGGCGGACGCGTGAATGATCGGAAGTATCCAGAGAGGGGGGTA 120  
DB 89 CTGGTGGCGATGGCGGACGCGTGAATGATCGGAAGTATCCAGAGAGGGGGGTA 148  
QY 121 ACCGATCGAAGAATGTGCTAATACCGCATATACCTTAAGAGAGAAAGCAGGGATCGAA 180  
DB 149 GCCCGGAGAAAGCGGATTAATACCGCATACGATCTACGATGAAAGCGGGGACCTTGC 208  
QY 181 GACCTTGGCGCTTTTGAAGCGCGCGATGCTGATTTAGTTAGTTGGTGGGTTAAAGCCTTAC 240  
DB 209 GGCCTGCGCTATAGG-GTTGGCGATGGCTGATTAGCTAGTTGGTGGGTTAAAGCCTTAC 267

QY 241 CAAGCGACGATCATAGTATTGGTCTGAGAGAGACCAACCACTAGGACTGAGACAG 300  
DB 268 CAGGCGACGATCATAGTATTGGTCTGAGAGAGACCAACCACTAGGACTGAGACAG 327  
QY 301 GCCCAGACTCTTACGCGAGGCGACGACGTGGGAAATTTTGGACATGGGCGCAAGCTGATC 360  
DB 328 GCCCAGACTCTTACGCGAGGCGACGACGTGGGAAATTTTGGACATGGGCGCAAGCTGATC 387  
QY 361 CAGCAATGCGCGTGTGATGAAGAAAGGCTTGGGTTGTAACTCTTTCAAGTCGAGAAAG 420  
DB 388 CAGCAATGCGCGTGTGATGAAGAAAGGCTTGGGTTGTAACTCTTTGTCCGAAAGA 447  
QY 421 AAGGTTACGCTTAATATCGTATGACCCATGACGCTATCGACAGAAAGACACCGGCTAAC 480  
DB 448 AATCCCTGGCTCTAATATCAAGTCGGGGATGACGCTACCGAAGAAATAGCACCGGCTAAC 507  
QY 481 TAGGTCCAGACAGCCCGGTAATACGTAGGGTGCAGACGCTTAATCGAAATTAATCTGGCGGT 540  
DB 508 TAGGTCCAGACAGCCCGGTAATACGTAGGGTGCAGACGCTTAATCGAAATTAATCTGGCGGT 567  
QY 541 AAGGCTGGCGAGGCGGCTTGTAACTCAAGATGTGAATCCCGGCGCTTAACCTGGGAAT 600  
DB 568 AAGCGTGGCGAGGCGGCTTGTAACTCAAGATGTGAATCCCGGCGCTTAACCTGGGAAT 627  
QY 601 TGGCTTTGAAACTACAAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTAGCAG 660  
DB 628 TGCATTTGTGATCTGGAGGCTAGAGTATGCGAGAGGGGGGTGAATTCACGCTGTAGCAG 687  
QY 661 TGAATGCGTAGAGATATGAAAGAACATGATGCGAAGGCGAGGCTCTGGGTTAACTACT 720  
DB 688 TGAATGCGTAGAGATATGAAAGAACATGATGCGAAGGCGAGGCTCTGGGCTAACTACT 747  
QY 721 GACGCTCAGACGAAAGGCTGGGAGCAACAGGATTAGTACCTGTAGTCCAGCGCC 780  
DB 748 GACGCTCAGACGAAAGGCTGGGAGCAACAGGATTAGTACCTGTAGTCCAGCGCC 807  
QY 781 CTAAACGATGCTCAACTAGTATTGGGCTTATTAAGCTTGTATACGAACTAACCGGTGA 840  
DB 808 CTAAACGATGCTCAACTAGTATTGGGCTTATTAAGCTTGTATACGAACTAACCGGTGA 867  
QY 841 AGTTGACGCGCTGGGAGATGCGGTGCGAAGATTTAAACTCAAGAAATTGACGGGAGCC 900  
DB 868 AGTTGACGCGCTGGGAGATGCGGTGCGAAGATTTAAACTCAAGAAATTGACGGGAGCC 927  
QY 901 GCAACAGCGGTGATTAATGATTAATTCGATGCAACGCGAAACCTTACTACCTT 960  
DB 928 GCAACAGCGGTGATTAATGATTAATTCGATGCAACGCGAAACCTTACTACCTT 987  
QY 961 GACATGTACGAATTTTCTAGAGATGATGATGCT--TCGGGAACGCTAACACAGGTG 1017  
DB 988 GACATGTGCGAATCTCTGCTGAGAGGTGGAGTGTCTGAAAGAAAGAACCGGCGCACAGTG 1047  
QY 1018 CTGCATGAGCTGTGCTGACGCTGTGCTGAGATGTTGGTTAAAGTCCGCAACGAGCGCA 1077  
DB 1048 CTGCATGAGCTGTGCTGACGCTGTGCTGAGATGTTGGTTAAAGTCCGCAACGAGCGCA 1107  
QY 1078 ACCCTTGCATTAATTAATGATCAATTTGTTGGGCACTTAAAGAGATCGGCGGTACAA 1137  
DB 1108 ACCCTTGTCTTAATGTTGCTAC-----GCAAGAGCACTCTAAGAGATCTGCGGTACAA 1162  
QY 1138 CCGAGGAAGATGGGATGATGCTCAAGTCTCTATGAGCCCTTATGGGTAGGGCTTACAGC 1197  
DB 1163 CCGAGGAAGATGGGATGATGCTCAAGTCTCTATGAGCCCTTATGGGTAGGGCTTACAGC 1222  
QY 1198 TAATCAATGCGCGCTTACAGAGGTTTCCCAACCCCGAGGGGAGCTTAATCTCAGAAAC 1257  
DB 1223 TAATCAATGCTGCGGAACAGAGGTTTCCCAACCCCGAGGGGAGCTTAATCTCAGAAAC 1282  
QY 1258 GCGTGTATGCTGCGAAGTCTGCACTGCACTCCGTAAGTGGGAATGCTGTATAT 1317  
DB 1283 CCATGCTATGCTGCGAATGCACTCTGCACTGCAAGTGAAGGCTGGAATGCTGTATAT 1342  
QY 1318 CCGGATCAGCATGTGCGGCTGAATACGTTCCCGGCTTGTATACACCGCCCGCTCACAC 1377

DB 1343 CGCGATCATGATCGCGCGGTGAATCGTCCCGGCTTGTACACCGCCCGTCACAC 1402  
QY 1378 CATTGGAGTGGGTTTACCGAAGCAGGTAGTCTAACCGTAAAGAGGCGCTTGCCACGG 1437  
DB 1403 CATGGAGTGGGTTTACCGAAGTGGCTAGTCTTACCGCAAGAGGACGCTACACCG 1462  
QY 1438 TGAGATTCACTGGGGTG 1457  
DB 1463 TAGATTCACTGGGGTG 1482  
RESULT 30  
ADB61689  
ID ADB61689 standard; DNA; 1535 BP.  
XX ADB61689;  
XX 04-DEC-2003 (first entry)  
XX 16S rRNA of Burkholderia cepacia DNA sequence.  
XX enriching mRNA, high quality bacterial mRNA; bacterial gene expression;  
XX poly-A tail; mRNA purification; oligo-dT capture;  
XX prokaryote mRNA purification; bridging oligonucleotide; targeting region;  
XX capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;  
XX eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;  
XX 28S eukaryotic RNA bridging oligonucleotide; ds.  
XX Burkholderia cepacia.  
XX WO2003054162-A2.  
XX 03-JUL-2003.  
XX 19-DEC-2002; 2002WO-US041014.  
XX 20-DEC-2001; 2001US-00029397.  
XX (AMBI-) AMBION INC.  
XX Murphy GL, Whitley JP;  
XX WPI; 2003-663255/62.  
XX  
XX Depleting or isolating targeted nucleic acids e.g. rRNA, involves using a  
XX bridging oligonucleotide comprising bridging region and a targeting  
XX region complementary to a targeted nucleic acid, and a capture  
XX oligonucleotide.  
PS Claim 4; Page 173-174; 208pp; English.  
XX  
XX This invention relates to a novel method for isolating, depleting or  
XX separating a targeted nucleic acid, such as rRNA, from a sample  
XX comprising targeted and non-targeted nucleic acids. It effects a way of  
XX enriching for non-targeted nucleic acids such as mRNAs. Isolating  
XX sufficient quantities of high quality bacterial mRNA is a demanding  
XX process which impedes analysis of bacterial gene expression in the  
XX presence of host cells. A small percentage of bacterial mRNAs may be poly  
XX -A tailed, but these are targeted for degradation and tend to be  
XX unstable. As a result, the commonly employed method for mRNA purification  
XX with eukaryotic cells, oligo-dT capture, is ineffective. The present  
XX invention provides an alternative, more suitable method for mRNA  
XX purification from prokaryotes. The method of the invention comprises the  
XX incubation of a sample with a bridging oligonucleotide (containing a  
XX targeting region) and subsequently incubating with a capture  
XX oligonucleotide allowing the isolation of the target from the sample. The  
XX method is useful for depleting or isolating targeted nucleic acid, for  
XX example rRNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S  
XX or 18S or eukaryotic 28S rRNA, from a sample. The rRNA sequence may  
XX comprise any one of 64 fully defined sequences as given in the  
XX specification. The present sequence is that of a DNA sequence which  
XX represents the sequence of 16S rRNA of Burkholderia cepacia related to

CC the invention.  
XX Sequence 1535 BP; 385 A; 355 C; 488 G; 307 T; 0 U; 0 Other;  
SQ  
Query Match 77.9%; Score 1135.2; DB 10; Length 1535;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 1278; Conservative 0; Mismatches 173; Indels 9; Gaps 3;  
QY 1 ATTGAACGCTGGCGGATGCTTTATCATGCAAGTGAACGGGAGACGGATGCTTGCAAT 60  
DB 29 ATTGAACGCTGGCGGATGCTTTATCATGCAAGTGAACGGGAGACGGATGCTTGCAAT 88  
QY 61 CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGGAACGATTCAGAAAGAGGGGGTAA 120  
DB 89 CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGGAACGATTCAGAAAGAGGGGGTAA 148  
QY 121 ACCGATCGAAAGATGTGCTAATATCCGCAATATCTCTAAGAGGAAGAACAGGGGATGMAA 180  
DB 149 GCCGGGGAAGAGCGGATTAATATCCGCAATATCTCTAAGAGGAAGAACAGGGGATGMAA 208  
QY 181 GACCTTGGCGCTTTTGAAGCGCGGATGTCTGATTAATTAAGTGGTGGGTTAAAGGCTTAC 240  
DB 209 GGCCTGGCGCTAATAGG-GTTGGGATGCTGATTAATTAAGTGGTGGGTTAAAGGCTTAC 267  
QY 241 CAAGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
DB 268 CAAGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 327  
QY 301 GCCCAGACTCTTACCGGAGGACAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCTTGATC 360  
DB 328 GCCCAGACTCTTACCGGAGGACAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCTTGATC 387  
QY 361 CAGCAATGCCGCGTGAAGTGAAGAAAGCCTTGGGTTGAAGCTCTTCACTGAGAGAGA 420  
DB 388 CAGCAATGCCGCGTGAAGTGAAGAAAGCCTTGGGTTGAAGCTCTTCACTGAGAGAGA 447  
QY 421 AAAGTTACGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
DB 448 AATTCCTGGCTTAAATCACTCGGGGATACGGTACCGGAATTAAGACCGGCTTAC 507  
QY 481 TACGTGCACAGCGCGGATTAATCACTGAGGATGCAAGCGTTAATCGGAATTAATCTGGGCGT 540  
DB 508 TACGTGCACAGCGCGGATTAATCACTGAGGATGCAAGCGTTAATCGGAATTAATCTGGGCGT 567  
QY 541 AAAGGATGCCAGCGCGCTTGAAGTCAATGATGAATCCCGGCTTAACTCGGGAAT 600  
DB 568 AAAGGATGCCAGCGCGCTTGAAGTCAATGATGAATCCCGGCTTAACTCGGGAAT 627  
QY 601 TGGCTTTGAAATCAAAAGCTAAGTGTGGCAGAGGAGTGAATTCATGTGTAGCAG 660  
DB 628 TGCATTGTGATCTGGCAGGCTAAGATGTGAGAGGGGGGTAGAAATTCACGTTAGACG 687  
QY 661 TGAATGCGTAGATATGGAAGAACATGATGCGAAGGACGCTCTGGGTTAAGT 720  
DB 688 TGAATGCGTAGATATGGAAGAACATGATGCGAAGGACGCTCTGGGTTAAGT 747  
QY 721 GACGCTCATGACGAAGAGCTGGGAGCAAAACAGATTAATATCCCTGTGTGTCAAGCC 780  
DB 748 GACGCTCATGACGAAGAGCTGGGAGCAAAACAGATTAATATCCCTGTGTGTCAAGCC 807  
QY 781 CTTAAACGATGCACTAGTCTTGGGCTTATTAAGCTGTGTGAAGAGCTTAAGCGCGTGA 840  
DB 808 CTTAAACGATGCACTAGTCTTGGGCTTATTAAGCTGTGTGAAGAGCTTAAGCGCGTGA 867  
QY 841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAATCAAAAGAAATTAAGCGGAGACC 900  
DB 868 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAATCAAAAGAAATTAAGCGGAGACC 927  
QY 901 GCACAGCGGTGATTAATGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 960  
DB 928 GCACAGCGGTGATTAATGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 987  
QY 961 GACATGTAGCAATTTTCTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1017



```

Db      599  TGCATTTTAACTACCGGGCTAGAGTGTCTCAGAGGAGGAGTTCCTCCGCTGTACAG 658
Qy      661  TGAATGCGTAGATATGGAAGACATGCATGSCAAGGACGCTCTGGGTTAACT 720
Db      659  TGAATGCGTAGATATGCGAGGACACCGATGCGAAGGACGCTCTGGGTTAACT 718
Qy      721  GAGCGTATGCAAGAAAGCGTGGGAGCAACAGATTAAGTACCTGTGATGTCACGCC 780
Db      719  GAGCGTATGCAAGAAAGCGTGGGAGCAACAGATTAAGTACCTGTGATGTCACGCC 778
Qy      781  CTAACGATGCACTAGTGTGTGGGCTTATTAGGCTTGGTAAAGCTAAAGCGCTGA 840
Db      779  CTAACGATGCACTAGTGTGTGGGCTTATTAGGCTTGGTAAAGCTAAAGCGCTGA 838
Qy      841  AGTTGACCGCTGGGGAGTACGGTCCGAAAGTAACTCAAGAGAAATTGACGGGAGCC 900
Db      839  AGTTGACCGCTGGGGAGTACGGTCCGAAAGTAACTCAAGAGAAATTGACGGGAGCC 898
Qy      901  GCACAAAGCGGTGATTAATGATGATTAATTCGATGCAACGCAAAACCTTACCTACCTT 960
Db      899  GCACAAAGCGGTGATTAATGATGATTAATTCGATGCAACGCAAAACCTTACCTACCTT 958
Qy      961  GACATGTAAGCAATTTTCTAGAGATGATTAAGTCT--TCGGAAAGCTTAAACAGGTG 1017
Db      959  GACATGTAAGCAATTTTCTAGAGATGATTAAGTCT--TCGGAAAGCTTAAACAGGTG 1018
Qy      1018  CTGATGAGCTTCTGACGCTGTCGAGATGTTGGGTTAAGTCCGCAAGAGAGCA 1077
Db      1019  CTGATGAGCTTCTGACGCTGTCGAGATGTTGGGTTAAGTCCGCAAGAGAGCA 1078
Qy      1078  ACCCTTGATTAATGTCATATTTGGTGGGCACTTAAATGAGACTGCGGTGACAA 1137
Db      1079  ACCCTTGATTAATGTCATATTTGGTGGGCACTTAAATGAGACTGCGGTGACAA 1133
Qy      1138  CCGAGAGAGGTGGGAGTACGTCAGTCTCATAGCCCTTAATGAGGTTGAGCTTACACG 1197
Db      1134  CCGAGAGAGGTGGGAGTACGTCAGTCTCATAGCCCTTAATGAGGTTGAGCTTACACG 1193
Qy      1198  TAAATCAATGCGGCTACAGAGGTTGCAACCGCGAGGGGAGCTAATCTCAGAAAGC 1257
Db      1194  TCAATCAATGCGGCTACAGAGGTTGCAACCGCGAGGGGAGCTAATCTCAGAAAGC 1253
Qy      1258  GCGTGTAGTCCGATCCGAGTCTGCAACTCGACTCCGTGAAGTCCGATCTAGTAAT 1317
Db      1254  CGATGTAGTCCGATCCGAGTCTGCAACTCGACTCCGTGAAGTCCGATCTAGTAAT 1313
Qy      1318  CGCGGATCAGATGTGCGGATGATACGTTCCGGGCTTGTGACACACCGCCGTCACAC 1377
Db      1314  CGCGGATCAGATGTGCGGATGATACGTTCCGGGCTTGTGACACACCGCCGTCACAC 1373
Qy      1378  CATGGAGTGGGTTTACCAAGAGAGTATACCGTAA--GGAAGGCGCTTGCCAC 1435
Db      1374  CATGGAGTGGGTTTACCAAGAGAGTATACCGTAA--GGAAGGCGCTTGCCAC 1433
Qy      1436  GGTGAGTTCACTGAGGGGTG 1457
Db      1434  GGTGAGTTCACTGAGGGGTG 1455

```

## RESULT 32

AAC86023 standard; cDNA; 1478 BP.

AAC86023;

29-AUG-2001 (fixed entry)

R. purpureus 16S rdna.

16S rdna; polyphosphate accumulating organism; PAO; probe: primer:  
detection; phosphorus; waste water; sludge; ss.

```

OS      Rhodocyclus purpureus.
XX
XX      W0200146459-A1.
PN
XX      28-JUN-2001.
PD
XX      28-DEC-2000; 2000MO-AU001611.
PF
XX      23-DEC-1999; 99AU-00004867.
PR
XX      (RCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.
PA
XX      Hogenholtz P, Crocetti GR, Tyson GW, Blackall LJ;
PI      WPI; 2001-408656/43.
DR
XX      Novel oligonucleotide probe or primer useful for detecting polyphosphate
PT      accumulating organism in a sample, comprises a sequence that is unique to
PT      16S rdna of polyphosphate accumulating organisms.
PS
XX      Claim 4; Fig 3; 54pp; English.
CC
XX      The sequences given in AAC86021-30 represent 16S rdna sequences from
CC      polyphosphate accumulating organisms (PAOs). Sequences which are unique
CC      to these 16S rdna sequences are used to create a probe or primer for
CC      detecting the relevant organisms. The primer/probe sequences are useful
CC      for detecting PAO cells in a sample, by treating cells in the sample to
CC      fix cellular contents, contacting fixed cells with the primer/probe which
CC      is labelled with a radiolabel, a reporter group or a hapten, under
CC      conditions which allow the probe to hybridize with 16S rdna within the
CC      fixed cell, removing unhybridized probe from the fixed cells, and
CC      detecting the labeled probe-RNA hybrid by fluorescence in situ
CC      hybridization. The primer/probe sequences are useful for identifying PAOs
CC      that are capable of biologically removing phosphorus from waste water.
CC      Rapid assessment of the presence of a number of PAOs in a waste water
CC      sample, can be done using the primer/ probe sequences. They allow quick
CC      and convenient assessment of whether a sludge or waste water sample
CC      includes PAOs and allows quantitation of PAO cells in samples
XX
XX      Sequence 1478 BP; 367 A; 345 C; 470 G; 289 T; 0 U; 7 Other;
SQ

```

Query Match 77.6%; Score 1130.6; DB 4; Length 1478;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 1269; Conservative 0; Mismatches 176; Indels 7; Gaps 3;

```

Qy      1  ATTGAAGCTGGCGGATGCTTTACATGCAAGTCGAAACGCGACGATGCTTGAT 60
Db      29  ATTGAAGCTGGCGGATGCTTTACATGCAAGTCGAAACGCGATGCTTGAT 86
Qy      61  CTGTGGCGAGTGGCGGACGGGTGATATCATCGGAACCTATCCAGAAAGGGGGTA 120
Db      87  CGCGGAACGAGTGGCGGAACGGGTGATATCATCGGAACCTATCCAGAAAGGGGGTA 146
Qy      121  ACGCATGAAAGATGTGTTATACCGCATATACCTTAAGAGGAAAGCAGGGAGATCGAAA 180
Db      147  ACGTAGCGAAAGTATACGTTATACCGCATATCTGTGACGAGAAAGCAGGGAGACTTCG 206
Qy      181  GACCTTGGCTTTTGGAGCGGCGCATGCTGATTAGCTAGTGTGGGTAAAGGCTTAC 240
Db      207  GGCCTTGGCTTTTGGAGTGGCGCATGCTGATTAGCTAGTGTGGGTAAAGGCTTAC 266
Qy      241  CAAGGCGAGATCAGTAGTGTGTTAGAGGACGACGACCACTGGGACTGAGACG 300
Db      267  CAAGGCGAGATCAGTAGTGTGTTAGAGGAGATGATCCGCCACTGGGACTGAGACG 326
Qy      301  GCCCAGATCTCTTAGGGAGGAGGAGCAAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360
Db      327  GCCCAGATCTCTTAGGGAGGAGGAGCAAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 386
Qy      361  CAGCAATGCGCGTGAAGAGAAAGGCTTGGGTTGTAAGCTCTTTCAGTCGAGAA 420
Db      387  CAGCAATGCGCGTGAAGAGAAAGGCTTGGGTTGTAAGCTCTTTCAGTCGAGAA 446

```

```

QY 421 AAAGTTACGGTAATTAATGTCGACCCATGACCGGTATCGA CAGAGAAGACCGGCTAAC 480
DB 447 AATCGGGTTTCTTAATACGGAACCCGGATGACGTTACCCAGAGAAGAGCA CCGGCTAAC 506
QY 481 TACGTGCACAGACCGCGGTAAATACGTAGGTGCAACCGTTAATCGAATTAATCTGGGCGT 540
DB 507 TAGGTGCACAGACCGCGGTAAATACGTAGGTGCAACCGTTAATCGAATTAATCTGGGCGT 566
QY 541 AAAGGTGGCCAGCGCGCTTGTAAATGTCAGATGTAATCCCGGGCTTAACTCGGGAAT 600
DB 567 AAAGCGTCCAGCGCGTGTGTAAAGACAGCGGAATCCCGGGCTCAACCTCGGGAAC 626
QY 601 TGGGTTGAAACTCAAAAGCTAGAGTGTGGCAGAGGAGGAGGATTCATGTTAGTACAG 660
DB 627 TGGGTTGTGACGTGACAGCTAGAGTACGCGAGAGGAGGAGGATTCATGTTAGTACAG 686
QY 661 TGAATGCTAGAGATATGGAAGAACATCGATGGCGAAGGACGCTCTGGGTTAACTACT 720
DB 687 TGAATGCTAGAGATATGGAAGAACATCGATGGCGAAGGACGCTCTGGGTTAACTACT 746
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAATGATCCTGCTGATTCACAGCC 780
DB 747 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAATGATCCTGCTGATTCACAGCC 806
QY 781 CTAAGCATGTCAACTAGTGTGTTG--GGCCTTATTAAGCTTGGTAAAGAACTAACGCGT 838
DB 807 CTAAGCATGTCAACTAGTGTGTTGTTGGTGGGTTAAACCAATTAAGTCCGTACAGCGT 866
QY 839 GAAATTGACCGCTGGGAGTACGCTGCAAGATTAATAACTCAAGAAATTAAGACGCGGAC 898
DB 867 GAAATTGACCGCTGGGAGTACGCGCGCAAGGTTAAATCTCAAGAAATTAAGACGCGGAG 926
QY 899 CCGCAACAAGCGGTGAGTATGTCGATTAATTCGATGCAAGCCGAAACCTTACCTAACCC 958
DB 927 CCGCAACAAGCGGTGAGTATGTCGATTAATTCGATGCAAGCCGAAACCTTACCTAACCC 986
QY 959 TTGACATGTAGCGCAATTTCTAGAGATAGATAGTATGTC---TTGCGGAACCTTACACAG 1015
DB 987 TTGACATGTAGCGCAATTTCTAGAGATAGATAGTATGTC---TTGCGGAACCTTACACAG 1046
QY 1016 TGTGTCATGTCGTGTCAGCTCGTGTCTGAGATGTTGGGTTAAGTCCGCAACGAGCG 1075
DB 1047 TGTGTCATGTCGTGTCAGCTCGTGTCTGAGATGTTGGGTTAAGTCCGCAACGAGCG 1106
QY 1076 CAACCTTGTCAATTAATTCGATCATTTGGTGGGCACTTAAATAGACCTCCGCTGACA 1135
DB 1107 CAACCTTGTCAATTAATTCGATCATTTGGTGGGCACTTAAATAGAACTCCGCTGACA 1166
QY 1136 AACCGAGAAAGGTGGGAGTACGTCAGATCTCATGAGCCCTTATGGGTAGGGCTTACA 1195
DB 1167 AACCGAGAAAGGTGGGAGTACGTCAGATCTCATGAGCCCTTATGGGTAGGGCTTACA 1226
QY 1196 CGTAATACAAATGCGCTACAGAGGCTTGCACACCCGCGAGGGGAGCTAATCTCAGAAA 1255
DB 1227 CGTAATACAAATGCGCTACAGAGGCTTGCACACCCGCGAGGGGAGCTAATCTCAGAAA 1286
QY 1256 GCGCGTGTGTCGTGTCGAGTCTGCAACTGATCTCGTGAAGTGGAAATGGCTAGTA 1315
DB 1287 GCGCGTGTGTCGTGTCGAGTCTGCAACTGATCTCGTGAAGTGGAAATGGCTAGTA 1346
QY 1316 ATGCGGATACAGATGTCGCGGTGAATACGTTCCCGGGCTTGTACACACCGCCGCTCAC 1375
DB 1347 ATGCGGATACAGATGTCGCGGTGAATACGTTCCCGGGCTTGTACACACCGCCGCTCAC 1406
QY 1376 ACCATGGAGTGGGTTTACCAAGAAAGAGTATCTAACCGTAAAGAGGCGCTTGGCAC 1435
DB 1407 ACCATGGAGTGGGTTTACCAAGAAAGTATTAAGCTTAACCGCAAGAGGCGCTTAACAC 1466
QY 1436 GGTGAGATTTCAT 1447
DB 1467 GGCAGCGTTTCGT 1478

```

```

RESULT 33
ID AD067895
AC AD067895;
DT 09-SEP-2004 (first entry)
XX
DE Acidovorax temperans partial 16S rRNA gene SEQ ID NO:2.
KW ds; gene; 16S rRNA; algae; algae-lysing bacteria; blue-green algae;
KW Anabaena cylindrica.
OS Acidovorax temperans.
PN KR2003075871-A.
XX
PD 26-SEP-2003.
XX
PF 21-MAR-2002; 2002KR-00015343.
PR 21-MAR-2002; 2002KR-00015343.
XX
PA (HANY-) HANYANG HAK WON CO LTD.
PI Bang SW, Choi JH, Han MS, Kim BH,
XX WPI; 2004-105157/11.
DR
XX
PT Algae-lysing bacteria Acidovorax temperans KFCC 11294 and method for
PT removing algae using the same microorganism.
XX
PS Example 3; SEQ ID NO 2; 1bp; Korean.
XX
CC The invention relates to a novel algae-lysing bacteria Acidovorax
CC temperans KFCC 11294 and a method for removing algae using the same
CC microorganism, thereby effectively solving environmental problems caused
CC by algae. An algae-lysing bacteria Acidovorax temperans KFCC 11294
CC containing 16S rDNA having the nucleotide sequence of AD067894 is
CC provided, wherein Acidovorax temperans KFCC 11294 is isolated from
CC Seokhon lake. A method for removing algae comprises culturing Acidovorax
CC temperans KFCC 11294 and applying the cultured medium to algae living in
CC water, where the cultured medium of Acidovorax temperans KFCC 11294 is
CC obtained by culturing Acidovorax temperans KFCC 11294 in BG-11 medium and
CC centrifuging the cultured medium. The algae is blue-green algae, and the
CC blue-green algae is Anabaena cylindrica. The present sequence represents
CC the Acidovorax temperans partial 16S rRNA gene.
XX
SQ Sequence 1509 BP; 389 A; 343 C; 477 G; 300 T; 0 U; 0 Other;
Query Match 77.5%; Score 1129.2; DB 12; Length 1509;
Best Local Similarity 87.3%; Pred. No. 0; Mismatches 173; Indels 12; Gaps 3;
Matches 1275; Conservative 0;
QY 1 ATGAAGCTGGCGGCGATCTTACATGCAAGTGAACGCGACGACGATGCTTGAT 60
DB 19 ATTGAACGCTGGCGGCGATCTTACATGCAAGTGAACGCGATGCTTGATGAT 75
QY 61 CTGTGTCGAGTGGCGGACGCGGTAGTAATGCATCGGAACGTAATCCAGAAAGGGGGTA 120
DB 76 -TGCTGACAGATGGCGGAAGCGGTGAGTAATACATCGGAACGTCGCCGATCGTGGGGGATA 134
QY 121 ACGCATCGAAAGATGCTAATACCGCATATCTCTAAGGAGGAAAGAGGGGATGAAA 180
DB 135 ACGAAACGAAAGCTTTGCTAATACCGCATTAAGTCTACGATGAAGAAAGAGGGGACCGCA 194
QY 181 GACCTTGGCGCTTTTGGAGCGCGGATGTCGATTAAGCTAGTTGTTGGGGTAAAGGCTTAC 240
DB 195 GGCCTTGGCGCGAAGCGAGCGCGGATGCGAATTAAGTGTGGTGAATTAAGGCTTAC 254
QY 241 CAAAGCGAGATCAATGTTGTTGTCGAGAGAGACAGACACGACCACTGGGATCGAGACAG 300
DB 255 CAAAGCGAGATCTGATGCTGTTGTCGAGAGAGAGACAGACCACTGGGATCGAGACAG 314

```

```
QY 301 GCCCAGACTCTCTACGGGAGGACGATGGGAAATTTTGAACAATGGCGCAGACTGATC 360
    |||||
Db 315 GCCCAGACTCTCTACGGGAGGACGATGGGAAATTTTGAACAATGGCGCAGACTGATC 374
    |||||
QY 361 CAGCAATGCGCGTGAAGAAAGGCTTTCGGGTTGTAAGCTTTTCAGTGAGAGA 420
    |||||
Db 375 CAGCAATGCGCGTGAAGAAAGGCTTTCGGGTTGTAAGCTTTTCAGTGAGAGA 434
    |||||
QY 421 AAAGTTACGGTAAATATATGTCACCATGACGGTATCGACAGAAGAAGCAGCGCTAC 480
    |||||
Db 435 AAAGACTCTGTTAATATCTGGGGTCCATGACGGTATCGTAAAGATTAAGCAGCGCTAC 494
    |||||
QY 481 TAGCTGCACAGCGCGGTAATATCTAGGGTGCACGCTTAATCGGAATTAATCTAGGCGT 540
    |||||
Db 495 TAGCTGCACAGCGCGGTAATATCTAGGGTGCACGCTTAATCGGAATTAATCTAGGCGT 554
    |||||
QY 541 AAAGGTCGCGAGCGGCTTGTAAAGTCAATGTGAATCCCGGGCTTAATCTGGGAT 600
    |||||
Db 555 AAAGCGTCGCGAGCGGCTTGTAAAGTCAATGTGAATCCCGGGCTTAATCTGGGAT 614
    |||||
QY 601 TCGCTTGAACCTTCAAGCTAGAGTGTGCGAGGGAGGTGGAATTCATGTGTAGAG 660
    |||||
Db 615 TGCATTTGTGACTGTATAGCTAGAGTACGCGAGGGAGGTGGAATTCATGTGTAGAG 674
    |||||
QY 661 TGAATGCGTAGATATGGAACAATCGATGCGCAAGCGCTCTCGGTTTAAGACT 720
    |||||
Db 675 TGAATGCGTAGATATGCGAGGAACACCGATGCGCAAGCGCTCTCGGCTGTAGACT 734
    |||||
QY 721 GACGCTATGACCAAAAGCGTGGGAGCAAAAGATTAATGATACCTGTGTATGATCAGCC 780
    |||||
Db 735 GACGCTATGACCAAAAGCGTGGGAGCAAAAGATTAATGATACCTGTGTATGATCAGCC 794
    |||||
QY 781 CTAACGATGTCAACTAGTTTGTGGCTTATTAAGCTGTGTAACGAAGCTTAAGCGCTGA 840
    |||||
Db 795 CTAACGATGTCAACTAGTTTGTGGCTTATTAAGCTGTGTAACGAAGCTTAAGCGCTGA 854
    |||||
QY 841 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAATACTCAAGGAATTAAGCGGAGCC 900
    |||||
Db 855 AGTTGACCGCTGGGAGTACGCTCGCAAGATTAATACTCAAGGAATTAAGCGGAGCC 914
    |||||
QY 901 GCACAAGCGGTGATTTATGTGATTAATTCATGACGCAAAACCTTAATCTTACCTT 960
    |||||
Db 915 GCACAAGCGGTGATTTATGTGATTAATTCATGACGCAAAACCTTAATCTTACCTT 974
    |||||
QY 961 GACATGTAGCAATTTTCTAAGATAGTATGTGT---TGGGAAACCTTAACAGGTG 1017
    |||||
Db 975 GACATGTAGCAATTTTCTAAGATAGTATGTGT---TGGGAAACCTTAACAGGTG 1034
    |||||
QY 1018 CTGATGGCTGTGCTCAGCTGCTGTGATGATGTTGGGTTAAGTCCGCAACGAGCGCA 1077
    |||||
Db 1035 CTGATGGCTGTGCTCAGCTGCTGTGATGATGTTGGGTTAAGTCCGCAACGAGCGCA 1094
    |||||
QY 1078 ACCCTTGTCAATTAATTTGCAATCTTTGTTGGCACTTTAATGAGACTCCGGTGACAA 1137
    |||||
Db 1095 ACCCTTGTCAATTAATTTGCTAGAAA-----GGGCACTCTAATGGGACTGCGGTGACAA 1149
    |||||
QY 1138 CCGAGAGAAAGTGGGAGTACGTCAGTCTCTATGCCCCCTTAATGGGTAGGGTTTACACG 1197
    |||||
Db 1150 CCGAGAGAAAGTGGGAGTACGTCAGTCTCTATGCCCCCTTAATGGGTAGGGTTTACACG 1209
    |||||
QY 1198 TAATATCATGCGCGGTAGACAGGGGTGCAACCGCGAGGGGGAGCTAATCTCAAAAGC 1257
    |||||
Db 1210 TCATACATGCGCTGTGATACAGGGGTGCAACCGCGAGGGGGAGCTAATCTCAAAAGC 1269
    |||||
QY 1258 GCGTGTAGTCCGGAATCGGAGTCTGCAACTCGACTCCGTGAAGTCGGAATGCTAGTAAT 1317
    |||||
Db 1270 CAGTGTAGTCCGGAATCGGAGTCTGCAACTCGACTCCGTGAAGTCGGAATGCTAGTAAT 1329
    |||||
QY 1318 CCGGATCAGCATGTGCGGGTAAATAGTTCCCGGGCTTTGTAGACACCGCCCGTCAAC 1377
    |||||
Db 1330 CCGGATCAGCATGTGCGGGTAAATAGTTCCCGGGCTTTGTAGACACCGCCCGTCAAC 1389
    |||||
```

```
QY 1378 CATGGAGTGGCTTTTCAACGAAGCAGGTAGTCTAACCGTAAGAGAGCGCTTGCCACG 1437
    |||||
Db 1390 CATGGAGCGGGTTCGCCAAGAGTAGTAGCTTAACGTAAGAGAGCGCTTAACACG 1449
    |||||
QY 1438 TGAGATTCACTGAGGAGTG 1457
    |||||
Db 1450 CAGGTTTCGTGACTGGGGTG 1469
    |||||

RESULT 34
ABZ69299
ID ABZ69299 standard; DNA; 1400 BP.
XX
AC ABZ69299;
XX
DT 11-AUG-2003 (first entry)
XX
DE J lividum 16s ribosomal RNA gene fragment #3.
XX
KW Osteoarthritis; antibacterial; Jantibobacterium; 16s RNA; gene;
XX osteopathic; antiarthritic; analgesic; anti-inflammatory; ds.
XX Jantibobacterium lividum.
XX OS
XX WO2002102384-A1.
XX
XX 27-DEC-2002.
XX
XX 17-JUN-2002; 2002WO-GB002771.
XX
XX 15-JUN-2001; 2001GB-00014672.
XX
XX (ORTH-) ORTHOGENICS AS.
XX PA
XX (GARD/) GARDNER R.
XX
XX EL-Gewely MR.
XX
XX WPI; 2003-175199/17.
XX
XX Use of an antibacterial agent in the manufacture of a medicament for
XX treating osteoarthritis.
XX
XX Example 1; Page 49-51; 89pp; English.
XX
XX The present invention relates to the use of an antibacterial agent in the
XX production of a treatment for osteoarthritis. The bacteria causing
XX osteoarthritis is Jantibobacterium lividum. The present sequence is a
XX fragment of the 16s RNA coding sequence from J. lividum shown in the
XX exemplification of the invention
XX
XX Sequence 1400 BP; 371 A; 309 C; 432 G; 288 T; 0 U; 0 Other;
XX

Query Match 77.4%; Score 1128.2; DB 8; Length 1400;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 143; Indels 10; Gaps 3;

QY 27 CATCAAGTCGCAACGCGACGACGATGCTTGATCTGTGTGCGAGTGGCGGACGGGTGAG 86
    |||||
Db 1 CATCAAGTCGCAACGCGACGACGGA--GCTTGTCTGTGTGCGAGTGGCGAAACGGGTGAG 58
    |||||
QY 87 TAATGATCGGAACGTATCCAGAAAGAGGGGGGTAAACGATGAAAGATGTGCTAATACG 146
    |||||
Db 59 TAATATATCGGAACGTATCCCTTAGAGTGGGGATTAAGTAAGAAATTAAGCTAATACG 118
    |||||
QY 147 CATATACCTTAAGAGGAAAGCAGGGGATCGAAAGACTTTCGCTTTTGAAGCGCGCAT 205
    |||||
Db 119 CATACGATCTTAAGAGTAAGTGGGGATTCGCAAGACTTATGCTGTGAGCGGCGAT 178
    |||||
QY 207 GTCTGATTAGCTAGTTGTGGGTAAAGCGCTTACCAAGCGACATCATAGTTGCTGTG 266
    |||||
Db 179 ATCTGATTAGTATGTTGGTAAAGCTTACCAAGGCATCATCATGATCATGATGCTGTG 238
    |||||
QY 267 AGAGAGCAGCAGCAGCAGCTGGAGCTAGACAGGCGCAAGCTCTACGGGAGGAGCAG 326
    |||||
```



Db	239	AAGAGACGACCGCCACTGGAACTGAGACACGGTCCAGACTCTTACGGAGGACGACG	298
Qy	327	TGGGGAAATTTTGGACATGGGGCGCAAGCCTGATCCAGCAATGCCGCTGAGTGAAGAAAG	386
Db	299	TGGGGAAATTTTGGACATGGGGGAAAGCCTGATCCAGCAATGCCGCTGAGTGAAGAAAG	358
Qy	387	CCTTGGGGTGTGTAAGCTTTTCAGTCGAGAAAGAAAGTTACGGTAATTAATCGTGACC	446
Db	359	CTTGGGGTGTGTAAGCTTTTGTCAAGGAAAGAAACGGTGAAGCTAATATCTTGTGT	418
Qy	447	CATGACGGTATGACAGAAAGAACGCCGCTTACTACGAGCCAGACGCCGCTTAATACG	506
Db	419	AATGACGGTACTGAAAGATTAAGCACCGCTTACTACGAGCCAGACGCCGCTTAATACG	478
Qy	507	TAGGGTGCAAAGCTTAATCGAATTACTGGGCGTAAAGGGTGCAGAGCGGCTTGTAA	566
Db	479	TAGGGTGCAAAGCTTAATCGAATTACTGGGCGTAAAGGGTGCAGAGCGGCTTGTAA	538
Qy	567	TCAGATGTGCAATCCCCGGGCTTAACTGGGAAATTCGCTTAAACTACAAAGCTAAGT	626
Db	539	TCGATGTGCAATCCCCGGGCTTAACTGGGAAATTCGATGAGCTGCAAGGCTAAGAT	598
Qy	627	GTGGCAGAGGGAGGTGGAATTCATGTGTGAGAGTGAATGGGTGAGATATGAAAGAC	686
Db	599	CTGGCAGAGGGGGGTGGAATTCACGTGTGAGAGTGAATGGGTGAGATATGTGAGAGAC	658
Qy	687	ATCGATGCGAAGGCAAGCCTCTGGGTTAACCTGACGCTCAACGAAACGTGGGGA	746
Db	659	ACCGATGCGAAGGCAAGCCTCTGGGTTCAAGATTGACGCTCAACGAAAGCGTGGGA	718
Qy	747	GCAACAGAGTTAAGTACCCTGGTGTAGTCCAGCCCTTAACGATGTCACTAGTTGTGG	806
Db	719	GCAACAGAGATTAGTACCCTGGTGTAGTCCAGCCCTTAACGATGTCTAGTTGTGG	778
Qy	807	CCTTATTAGGCTTGGTTAAAGAAAGCTAAGCGGTGAAGTTGACCGCTGGGGAGTACGTG	866
Db	779	TCTTATTAGCTTGGTTAAAGCAAGCTAAGCGGTGAAGTTGACCGCTGGGGAGTACGTG	838
Qy	867	CAGATTTAAACTCAAAGGAATTGACGGGGAACCCGCAACAGCGGTGATATGTGGATT	926
Db	839	CAGATTTAAACTCAAAGGAATTGACGGGGAACCCGCAACAGCGGTGATATGTGGATT	898
Qy	927	ATTGATGCCAACGGGAAAAACCTTACCTTACCTTGAACATGTAGGAAATTTCTGAGATA	986
Db	899	ATTGATGCCAACGGGAAAAACCTTACCTTACCTTGAACATGTGTAAGTCTGAGAGATT	958
Qy	987	GATTAGTCT--TCGGGACGCTTAACAAGGTGCTGATGGCTGTCTCACTGCTGTCTC	1043
Db	959	GAGAGTCTCTGGAAGAGAACCAAGTACACAGGTGTGATGGCTGTCTCACTGCTGTCTC	1018
Qy	1044	GTGAGATTTGGGTTAAGTCCCGCAACGAGGCAACCTTGTCAATTAATTGGCATCATTT	1103
Db	1019	GTGAGATTTGGGTTAAGTCCCGCAACGAGGCAACCTTGTCAATTAATTGGCATCATTT	1073
Qy	1104	GATTGGGCACTTTAATGAGACTGCGCGGTGACAAACCGAGAGAAAGGTGGGATGACGTCA	1163
Db	1074	GAAAGGCACTTTAATGAGACTGCGCGGTGACAAACCGAGAGAAAGGTGGGATGACGTCA	1133
Qy	1164	GTCTCATAGGCCCTTATGGGTAGGGCTTCAACGTTAATCAATGGCGGCTACAGAGGTT	1223
Db	1134	GTCTCATAGGCCCTTATGGGTAGGGCTTCAACGTTAATCAATGGTACATACAGAGGCC	1193
Qy	1224	GCCAACCCGCGAGGGGGAGCTTAATTCAGAAAGCGCGTGTGATGTCCGATGTGAGTCTGC	1283
Db	1194	GCCAACCCGCGAGGGGGAGCTTAATTCGAGAAAGTGTATGTATGTCCGATGTGATGTCTGC	1253
Qy	1284	AACCTGCACTCGGTGAAGTCCGGAATGTGCTAGTAATGCGGAGATCAGCATGTCCGGGTAATA	1343
Db	1254	AACCTGCACTCGGTGAAGTGTGTAATGTGCTAGTAATGCGGATCAGCATGTCCGGGTAATA	1313
Qy	1344	CGTTCCGGGCTTGTETACACACCGCCGTCACACCATGGAGGTGGTTTCCAGAAAGCA	1403

Db	1314	CGTCCCGGCGCTTGTATACACACCGCCCGCTACACCAATGGAGCGGCTTTACCGAAGTA	1373
Qy	1404	GGTAGCTTAACCGTAAGAGGCGCCTT	1430
Db	1374	GGTAGCTTAACCGCAAGAGGCGCCTT	1400
RESULT 35			
ID	ADB61688		
ID	ADB61688	standard; DNA; 1464 BP.	
XX	AC	ADB61688;	
XX	XX		
DT	04-DEC-2003	(first entry)	
XX	DE		
DE	16S rRNA of Bordetella pertussis DNA sequence.		
KW		enriching mRNA; high quality bacterial mRNA; bacterial gene expression;	
KW		poly-A tail; mRNA purification; oligo-dT capture;	
KW		prokaryote mRNA purification; bridging oligonucleotide; targeting region;	
KW		capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;	
KW		eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;	
KW		28S eukaryotic RNA bridging oligonucleotide; ds.	
XX			
OS		Bordetella pertussis.	
XX			
PN	WO2003054162-A2.		
PD	03-JUL-2003.		
XX			
PF	19-DEC-2002; 2002WO-US041014.		
PR	20-DEC-2001; 2001US-00029397.		
XX			
PA	(AMBI-) AMBION INC.		
PL	Murphy GL, Whitley JP;		
DR	WPI; 2003-663255/62.		
PT	Depleting or isolating targeted nucleic acids e.g. rRNA, involves using a		
PT	bridging oligonucleotide comprising bridging region and a targeting		
PT	region complementary to a targeted nucleic acid, and a capture		
PT	oligonucleotide.		
XX			
PS	Claim 4; Page 173; 208pp; English.		
XX			
CC	This invention relates to a novel method for isolating, depleting or		
CC	separating a targeted nucleic acid, such as rRNA, from a sample		
CC	comprising targeted and non-targeted nucleic acids. It effects a way of		
CC	enriching for non-targeted nucleic acids such as mRNAs. Isolating		
CC	sufficient quantities of high quality bacterial mRNA is a demanding		
CC	process which impedes analysis of bacterial gene expression in the		
CC	presence of host cells. A small percentage of bacterial mRNAs may be poly		
CC	-A tailed, but these are targeted for degradation and tend to be		
CC	unstable. As a result, the commonly employed method for mRNA purification		
CC	with eukaryotic cells, oligo-dT capture, is ineffective. The present		
CC	invention provides an alternative, more suitable method for mRNA		
CC	purification from prokaryotes. The method of the invention comprises the		
CC	incubation of a sample with a bridging oligonucleotide (containing a		
CC	targeting region) and subsequently incubating with a capture		
CC	oligonucleotide allowing the isolation of the target from the sample. The		
CC	method is useful for depleting or isolating targeted nucleic acid, for		
CC	example rRNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S		
CC	or 18S or eukaryotic 28S rRNA, from a sample. The rRNA sequence may		
CC	comprise any one of 64 fully defined sequences as given in the		
CC	specification. The present sequence is that of a DNA sequence which		
CC	represents the sequence of 16S rRNA of Bordetella pertussis related to		
CC	the invention.		
XX			
SQ	Sequence 1464 BP; 361 A; 344 C; 468 G; 285 T; 0 U; 6 Other;		
Query March	77.4%; Score 1128.2; DB 10; Length 1464;		



Best Local Similarity 87.6%; Pred. No. 0;  
Matches 1266; Conservative 0; Mismatches 169; Indels 10; Gaps 3;

```
QY 1 ATTGAAGCGTGGCGGAGCTTTACATGCAAGTCGAAAGCGGACGAGTGGTTCAT 60
Db 27 ATTGAAGCGTGGCGGAGCTTTACATGCAAGTCGAAAGCGGACGAGTGGTTCAT 60
QY 61 CTGGTGGCGAGTGGCGGAGCTTTACATGCAAGTCGAAAGCGGACGAGTGGTTCAT 120
Db 85 CTGGTGGCGAGTGGCGGAGCTTTACATGCAAGTCGAAAGCGGACGAGTGGTTCAT 144
QY 121 ACCGATGAAGATGCTGCTAATACCGATTAATCTGTAAGAGGAAAGCAGGAGGAT 180
Db 145 ACTACGAGAAAGCTGATCTAATACCGATTAATCTGTAAGAGGAAAGCAGGAGGAT 204
QY 181 GACCTTGGCTTTTGGAGCGGCGGATGCTGTAAGTGGTGGGAGTAAAGCGCTTAC 240
Db 205 GGCCTCCGACTATTTGAGCGGCGGATGCTGTAAGTGGTGGGAGTAAAGCGCTTAC 264
QY 241 CAAGGCGACGATCAGTGGTGGTGGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300
Db 265 CAAGGCGACGATCAGTGGTGGTGGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 324
QY 301 GCCGAGACTCTTACGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 360
Db 325 GCCGAGACTCTTACGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 384
QY 361 CAGCAATGCCGCGTGGAGAGAGGCGCTTGGGTTGTAAGCTTTTCACTGAGAGA 420
Db 385 CAGCAATGCCGCGTGGAGAGAGGCGCTTGGGTTGTAAGCTTTTCACTGAGAGA 444
QY 421 AAAGGTTACGGTAATATCGTGACCCATGACGCGATGACAGAGAGAGCAGCGCTTAC 480
Db 445 AAAGGTTACGGTAATATCGTGACCCATGACGCGATGACAGAGAGAGCAGCGCTTAC 504
QY 481 TAGGTGCCAGAGCGCGGTAATATCGTGACCGATGACAGAGAGAGCAGCGCTTAC 540
Db 505 TAGGTGCCAGAGCGCGGTAATATCGTGACCGATGACAGAGAGAGCAGCGCTTAC 564
QY 541 AAAGGTTACGGTAATATCGTGACCCATGACGCGATGACAGAGAGAGCAGCGCTTAC 600
Db 565 AAAGGTTACGGTAATATCGTGACCCATGACGCGATGACAGAGAGAGCAGCGCTTAC 624
QY 601 TGCGTTGAAACTACAAAGCTAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 625 TGCGTTGAAACTACAAAGCTAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 684
QY 661 TGAATGGGTAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 685 TGAATGGGTAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 744
QY 721 GACGCTCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 745 GACGCTCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804
QY 781 CTAAACGATGCACTAGTGGTGGGCTTATTAAGGCTTGGTAAGAGAGAGAGAGAGAG 840
Db 805 CTAAACGATGCACTAGTGGTGGGCTTATTAAGGCTTGGTAAGAGAGAGAGAGAGAG 864
QY 841 AGTTGACCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 865 AGTTGACCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924
QY 901 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 925 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 984
QY 961 GACATGTCTGGAATCCCGAAGAGATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
Db 985 GACATGTCTGGAATCCCGAAGAGATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1044
QY 1018 CTGACATGCTGTCTGAGCTGTCTGAGAGATTTGGGTTAAGTCCCGCAAGAGAGAG 1077
```

```
Db 1045 CTGACATGCGTGTCTGAGCTGTCTGAGAGATTTGGGTTAAGTCCCGCAAGAGCGCA 1104
QY 1078 ACCCTTGTCTAATTTGCTTATTTGCTTGGTGGGCACTTAAATGAGACTGCGGAGCAAA 1137
Db 1105 ACCCTTGTCTAATTTGCTTATTTGCTTGGTGGGCACTTAAATGAGACTGCGGAGCAAA 1159
QY 1138 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
Db 1160 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1219
QY 1198 TAAATCAATGCGCGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
Db 1220 TCAATCAATGCGCGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1279
QY 1258 GCGCTGTAAGTCCGAGTGGAGAGTCTGCAATCTGCACTCCGTGAATCCGGAATCCGTA 1317
Db 1280 GCGCTGTAAGTCCGAGTGGAGAGTCTGCAATCTGCACTCCGTGAATCCGGAATCCGTA 1339
QY 1318 CCGGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
Db 1340 CCGGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1399
QY 1378 CATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
Db 1400 CATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1459
QY 1438 TGAGA 1442
Db 1460 TAGGA 1464

RESULT 36
ADBE1077
ID ADBE1077 standard; DNA; 1481 BP.
XX
AC ADEB1077;
XX
DT 29-JAN-2004 (first entry)
XX
DE Acidovorax avenae subsp. citrulli 16S ribosomal RNA gene SEQ ID NO:1.
XX
KW 16S-23S spacer DNA; bacterial; Acidovorax avenae subsp. citrulli;
KW Xanthomonas cucurbitae; Erwinia tracheiphila; bacterial pathogen;
KW 16S ribosomal RNA; gene; ds.
XX
OS Acidovorax avenae subsp. citrulli.
XX
PN WO2003082079-A2.
XX
PD 09-OCT-2003.
XX
PF 25-MAR-2003; 2003WO-US009618.
XX
PR 25-MAR-2002; 2002US-0367628P.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Beck JY, Barnett CJ, Yarnall WS, Zeilouni L;
XX
DR WPI: 2003-902900/82.
XX
PT New nucleic acid encoding a 16S-23S spacer DNA sequences for bacterial
XX species Acidovorax avenae subsp. citrulli, Xanthomonas cucurbitae and
XX Erwinia tracheiphila, useful in identifying A. avenae subsp. citrulli
XX in melons.
XX
PS Example 8; SEQ ID NO 1; 50pp; English.
XX
CC The present invention describes a nucleic acid molecule encoding a 16S-
XX 23S spacer DNA sequence for the bacterial species Acidovorax avenae
XX subsp. citrulli or avenae, Xanthomonas cucurbitae and Erwinia
XX tracheiphila. Also described: (1) a pair of oligonucleotide primers,
XX where at least one primer comprises 15-18, 15-22 or 20-25 bp; (2)
```

CC detecting bacterial pathogen; (3) a diagnostic kit used in detecting the  
CC bacterial pathogen; (4) a polypeptide comprising the amino acid sequence  
CC (S1, see AB11075); (5) an antibody that reacts with a polypeptide having  
CC the N-terminal amino acid sequence of the polypeptide; and (6) an  
CC immunoassay for detecting Acidovorax avenae subspecies citrullii that uses  
CC the antibody. The nucleic acid is useful in diagnostic assays for  
CC identifying Acidovorax avenae subspecies citrullii in melons. The present  
CC sequence represents an Acidovorax avenae subspecies citrullii 16S  
CC ribosomal RNA gene, which is used in the exemplification of the present  
CC invention.

XX  
SQ Sequence 1481 BP; 378 A; 341 C; 470 G; 292 T; 0 U; 0 Other;

Query Match 77.4%; Score 1127.6; DB 10; Length 1481;

Best Local Similarity 87.3%; Pred. No. 0;

Matches 1274; Conservative 0; Mismatches 174; Indels 12; Gaps 3;

Qy 1 ATTGAAGCTGGCGGCGATGCTTTACATGCAAGTCGAAAGCGACAGCATGCTTGCAT 60  
Db 1 ATTGAAGCTGGCGGCGATGCTTTACATGCAAGTCGAAAGCGATGCTTCCGA--- 57  
Qy 61 CTGGTGGCGAGTGGCGGACGGGTGATGATCGAAGCTATCCAGAAAGGGGGGTA 120  
Db 58 -TCTGACGAGTGGCGAAAGGGGTGATGATCGAAGCTGCGATCGTGGGGGATTA 116  
Qy 121 ACCGATCGAAGATGCTAATACCGCATATACCTTAAGGAGGAAAGCAGGGGATCGAA 180  
Db 117 ACAGAGGAAAGCTTGTCTAATACCGCATAGATCTATGATGAAACAGAGGACGTTAA 176  
Qy 181 GACCTTGCGCTTTTGGAGCGGCGCATGCTGATTAAGTGGTGGGGTAAAGGCTTAC 240  
Db 177 GGGCTTTCGGGACGAGCGGCGGCGATGAGATTAGTAGTGGTGGGGTAAAGGCTTAC 236  
Qy 241 CAAGGCGACATCATGATGTTGTTGTGAGAGACGACCGACACTGGGACTGAGACAG 300  
Db 237 CAAGCTACGATCTGTAGTGTCTGAGAGACGACCGACACTGGGACTGAGACAG 296  
Qy 301 GCCCAAGCTCTTACGGGAGGACGAGTGGGGAATTTTGAACAATGGGCGGAAGCTGATC 360  
Db 297 GCCCAAGCTCTTACGGGAGGACGAGTGGGGAATTTTGAACAATGGGCGGACTGATC 356  
Qy 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCGGGTTTAAAGCTCTTTCAGTGGAGAGA 420  
Db 357 CAGCCATGCCGCGTGAAGTGAAGAGGCTTCGGGTTTAAAGCTCTTTCAGTGGAGAGA 416  
Qy 421 AAAGTTACGGTAAATATATCTGACCATACGGTATCGACAGAGAGACCGGCTTAC 480  
Db 417 AAAGCTTCTTCTAATTAAGGGGGTCTATGACGGTACCTTAAGAAATAGCACCGGCTTAC 476  
Qy 481 TAGGTGCACACACCGCGGTAAATACGTAAGGTGCAACCGTTAATCGGAATTAATCTGGGCGCT 540  
Db 477 TAGGTGCACACACCGCGGTAAATACGTAAGGTGCAACCGTTAATCGGAATTAATCTGGGCGCT 536  
Qy 541 AAAGGTCGCGACGCGGCTTGTAAAGTCAAGTGAATCCCGGGCTTAACCTGGGAAT 600  
Db 537 AAAGCGTGGCGACGCGGTGATGTAAGCAATGTAATCCCGGGCTTCAACCTGGGAAC 596  
Qy 601 TGGGTTTGAATCAAAAGCTAAGTGTGCGAAGGAGGTGGAATTCATGTGTACAG 660  
Db 597 TGCATTTGTGACTGATCGCTGAGTACGGCAGAGGGGGATGGAATTCGCCGTGTACAG 656  
Qy 661 TGAATGCGGTAAATATGGAAGAAATCGATGGCGGAAGGACGCTTCGGGTTAACAT 720  
Db 657 TGAATGCGGTAAATATGGAAGAAATCGATGGCGGAAGGACGCTTCGGGTTAACAT 716  
Qy 721 GACGCTCATGACGAAAGCTGGGGAGCAAAAGATTAATACCTGTATGTCACGCC 780  
Db 717 GACGCTCATGACGAAAGCTGGGGAGCAAAAGATTAATACCTGTATGTCACGCC 776  
Qy 781 CTAAACGATGTCAACTAGTTGTGGGCTTATTAAGCTTGATGCAAGACTTAACGCGTGA 840  
Db 777 CTAAACGATGTCAACTAGTTGTGGGCTTCACTGACTAGTAACGAAGCTTAACGCGCTGA 836

Qy 841 AGTTACCGCCTGGGGAGTACGGTCGCAAGATTAAACTCAAGAAATTGACGGGACCC 900  
Db 837 AGTTACCGCCTGGGGAGTACGGTCGCAAGATTAAACTCAAGAAATTGACGGGACCC 896  
Qy 901 GCAAGAGCGGTGATTAATGATTAATTAATGATGCAAGCGGAAACCTTACTTACCTT 960  
Db 897 GCACAGCGGTGATGATGATGATTAATGATGATGATGATGATGATGATGATGATGAT 956  
Qy 961 GACATTAACGAAATTTTCTAGAGATGATTAATGATGATGATGATGATGATGATGATGAT 1017  
Db 957 GACATTAACGAAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1016  
Qy 1018 CTGATGAGCTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077  
Db 1017 CTGATGAGCTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076  
Qy 1078 ACCCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
Db 1077 ACCCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131  
Qy 1138 CCGAGAGAGTGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
Db 1132 CCGAGAGAGTGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191  
Qy 1198 TAATCAATGAGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAAAGC 1257  
Db 1192 TCATCAATGAGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAAAGC 1251  
Qy 1258 GCGGTGATGCTGGATGCGAGTGTGCAACTGCACTCGTGAAGTGGGAATTCGTATTAAT 1317  
Db 1252 GCGGTGATGCTGGATGCGAGTGTGCAACTGCACTCGTGAAGTGGGAATTCGTATTAAT 1311  
Qy 1318 CCGGATGACAGATGTCGCGGTGAATTCGTTCCGGGCTTTGTAACACAGCGCGGTACAC 1377  
Db 1312 CCGGATGACAGATGTCGCGGTGAATTCGTTCCGGGCTTTGTAACACAGCGCGGTACAC 1371  
Qy 1378 CATGGAGTGGGTTTACAGAGAGCAGTATGTTAAACCTTAAGAGAGGGCGCTTGCACCG 1437  
Db 1372 CATGGAGTGGGTTTACAGAGAGCAGTATGTTAAACCTTAAGAGAGGGCGCTTGCACCG 1431  
Qy 1438 TGAGATTCAATGACTGGGGTG 1457  
Db 1432 CAGGGTTCGTGACTGGGGTG 1451

RESULT 37  
AD067894  
ID AD067894 standard; DNA; 1509 BP.  
XX  
AC AD067894;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
DE Acidovorax temperans partial 16S rRNA gene SEQ ID NO:1.  
XX  
XX ds: gene: 16S rRNA; algae: algae-Lysing bacteria; blue-green algae;  
XX Anabaena cylindrica.  
XX  
OS Acidovorax temperans; Kfoc 11294.  
XX  
PN KR2003075871-A.  
XX  
PD 26-SEP-2003.  
XX  
XX 21-MAR-2002; 2002KR-00015343.  
XX  
XX 21-MAR-2002; 2002KR-00015343.  
XX  
XX (HANY-) HANYANG HAK MON CO LTD.  
XX  
XX Bang SW, Choi JH, Han MS, Kim BH;  
XX  
XX WPI; 2004-105157/11.  
DR

XX Algae-lysing bacteria Acidovorax temperans KFC11294 and method for  
PT removing algae using the same microorganism.

XX Claim 1; SEQ ID NO 1; 1bp; Korean.

XX The invention relates to a novel algae-lysing bacteria Acidovorax  
CC temperans KFC11294 and a method for removing algae using the same  
CC microorganism, thereby effectively solving environmental problems caused  
CC by algae. An algae-lysing bacteria Acidovorax temperans KFC11294  
CC containing 16S rDNA having the nucleotide sequence of A0067894 is  
CC provided, wherein Acidovorax temperans KFC11294 is isolated from  
CC Sookchon lake. A method for removing algae comprises culturing Acidovorax  
CC temperans KFC11294 and applying the cultured medium to algae living in  
CC water, where the cultured medium of Acidovorax temperans KFC11294 is  
CC obtained by culturing Acidovorax temperans KFC11294 in BG-11 medium and  
CC centrifuging the cultured medium. The algae is blue-green algae, and the  
CC blue-green algae is Anabaena cylindrica. The present sequence represents  
CC the Acidovorax temperans KFC11294 16S rRNA gene.

XX Sequence 1509 BP; 386 A; 347 C; 477 G; 299 T; 0 U; 0 Other;

Query Match 77.4%; Score 1127.6; DB 12; Length 1509;  
Best Local Similarity 87.3%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 174; Indels 12; Gaps 3;

QY 1 ATTGAAGCGTGGCGGAGTCTTACATGCAAGTCGAAACGGCAGACGATGCTTCAT 60  
DB 19 ATTGAAGCGTGGCGGAGTCTTACATGCAAGTCGAAACGGTTCGGA--- 75  
QY 61 CTGGTGGCGAGTGGCGGAGTGAATGATCATCGAAACGTATCCAGAAAGGGGGGTA 120  
DB 76 -TGTGAGAGTGGCGGAGTGAATGATCATCGAAACGTATCCAGAAAGGGGGGTA 134  
QY 121 ACGCATGAAAGATGTGCTAATACCGCATATCTTAAGAGAAAGAGGGGATCGAA 180  
DB 135 ACGAAGCGAAAGCTTGTCTAATACCGCATATCTTAAGAGAAAGAGGGGATCGAA 194  
QY 181 GACCTTGGCTTTTGGAGCGGCGGATGCTGATTAGCTAGTGGGGTAAAGGCTTAC 240  
DB 195 GGGCTTGGCGGAGCGGCGGCGGATGCTGATTAGCTAGTGGGGTAAAGGCTTAC 254  
QY 241 CAAGGCGAGATCACTAGTATGCTGAGAGGAGCGACGACCACTGGGACTGAGACAG 300  
DB 255 CAAGCGAGATCTGTATCTGCTGTGAGAGGAGCGACGACCACTGGGACTGAGACAG 314  
QY 301 GCCCAGACTCTCTAAGGAGGAGCGAGTGGGAAATTTTGAACAATGGGGCGAAGCTGATC 360  
DB 315 GCCCAGACTCTCTAAGGAGGAGCGAGTGGGAAATTTTGAACAATGGGGCGAAGCTGATC 374  
QY 361 CAGCAATGCCCGTGAAGTGAAGGAGGCGGCTTGAAGCTCTTCACTGAGAGAGA 420  
DB 375 CAGCATGCCCGTGAAGTGAAGGAGGCGGCTTGAAGCTCTTCACTGAGAGAGA 434  
QY 421 AAAGGTTACGGTAAATATCTGACCATGACGATATGACAGAAAGACCGGCTAAC 480  
DB 435 AAAGGTTACGGTAAATATCTGACCATGACGATATGACAGAAAGACCGGCTAAC 494  
QY 481 TACGTGCGAGCGCGGTAATACGTAGGTCGAAAGCTTAATGGAATTACTGGGGCT 540  
DB 495 TACGTGCGAGCGCGGTAATACGTAGGTCGAAAGCTTAATGGAATTACTGGGGCT 554  
QY 541 AAAGGTCGCGAGCGGCTTGTAGTCAATGTAATCCCGGGCTTAACCTGGGAAT 600  
DB 555 AAAGGTCGCGAGCGGCTTGTAGTCAATGTAATCCCGGGCTTAACCTGGGAAT 614  
QY 601 TCGCTTTGAACTAACAACTAGATGTGGCAGAGGAGGTGAATTCATGATGAGCAG 660  
DB 615 TCGATTTGAGCTATAGTATGATGAGTACGAGAGGAGGTGAATTCATGATGAGCAG 674  
QY 661 TGAATTCGTAGATATGAGAAACATGATGGGAGAGCGAGCTCTGGGTTAACTACT 720  
DB 675 TGAATTCGTAGATATGAGAAACATGATGGGAGAGCGAGCTCTGGGTTAACTACT 734

QY 721 GACCTTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGGTATGCCAGCC 780  
DB 735 GACCTTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGGTATGCCAGCC 794  
QY 781 CTAACGATGTCAACTAGTGTGGGCTTATATAGGCTTGTATACAGAACTTAACGGGTGA 840  
DB 795 CTAACGATGTCAACTAGTGTGGGCTTATATAGGCTTGTATACAGAACTTAACGGGTGA 854  
QY 841 AGTTGACCGCTGGGAGTACGCGTCCGAAATTTAAACTCAAGAAATTGACGGGAGCC 900  
DB 855 AGTTGACCGCTGGGAGTACGCGTCCGAAATTTGAAATTCGAAATTTGACGGGAGCC 914  
QY 901 GCAACAGCGGTGATATATGATGATTAATTCATGCAACGGAACCTTAACCTTACCTT 960  
DB 915 GCAACAGCGGTGATATATGATGATTAATTCATGCAACGGAACCTTAACCTTACCTT 974  
QY 961 GACATGATGGAATTTTCTAGATATGATGATGCT---TCGGAAACGCTTAACAGAGTG 1017  
DB 975 GACATGATGGAATTTTCTAGATATGATGATGCT---TCGGAAACGCTTAACAGAGTG 1034  
QY 1018 CTGATGCTGTGCTGACGCTGTGTGCTGATGATGTTGGTTAACTCCCGCAAGAGCGCA 1077  
DB 1035 CTGATGCTGTGCTGACGCTGTGTGCTGATGATGTTGGTTAACTCCCGCAAGAGCGCA 1094  
QY 1078 ACCCTTGCATTTAATGCTCATCATTTGGTGGGCACTTATATGAGACTGCGGAGCAAA 1137  
DB 1095 ACCCTTGCATTTAATGCTCATCATTTGGTGGGCACTTATATGAGACTGCGGAGCAAA 1149  
QY 1138 CCGGAGGAAGTGGGAGTGAAGTCAAGTCTCATAGCCCTTAATGGTATGGGCTTACAGACG 1197  
DB 1150 CCGGAGGAAGTGGGAGTGAAGTCAAGTCTCATAGCCCTTAATGGTATGGGCTTACAGACG 1209  
QY 1198 TAATCAATGCGCGCTGACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAAAAGC 1257  
DB 1210 TAATCAATGCGCGCTGACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAAAAGC 1269  
QY 1258 GCGTGTATGTCGAGATGAGAGTCTGCAATCTGACTCCGTAAAGTCGAAATCGCTAGTAAT 1317  
DB 1270 GCGTGTATGTCGAGATGAGAGTCTGCAATCTGACTCCGTAAAGTCGAAATCGCTAGTAAT 1329  
QY 1318 CGCGGATCAGATGTCGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1377  
DB 1330 CGCGGATCAGATGTCGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1389  
QY 1378 CATGGAAGCGGCTTCTGCGAAGATGATGATGATGATGATGATGATGATGATGATGATG 1437  
DB 1390 CATGGAAGCGGCTTCTGCGAAGATGATGATGATGATGATGATGATGATGATGATGATG 1449  
QY 1438 TGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1457  
DB 1450 TGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1469

RESULT 38  
ADCS3929  
ID ADCS3929 standard; DNA; 1463 BP.  
XX ADCS3929;  
AC 18-DEC-2003 (first entry)  
XX  
DE Phenyl hydroxylase gene fragment #SEQ ID 1.  
XX  
XX Gene fixing; ecosystem; phenyl hydroxylase gene; phenol decomposition;  
KW sludge; purification; ds.  
XX  
OS Comamonas sp.  
XX JP2003144165-A.  
XX  
XX  
XX  
PD 20-MAY-2003.

PF 12-NOV-2001; 2001JP-00346473.  
XX  
XX 12-NOV-2001; 2001JP-00346473.  
XX  
XX (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.  
XX WPI; 2003-818164/77.  
XX  
PT Fixing useful genes e.g. phenol hydroxylase to micro organism ecosystems  
PT e.g. phenol decomposition activated sludge, involves isolating dominant  
PT strain in ecosystem, introducing a useful gene and returning the strain  
PT to the ecosystem.  
XX  
PS Example 2; SEQ ID NO 1; 20pp; Japanese.  
XX  
XX The invention relates to a method for fixing a useful gene in  
XX microorganisms ecosystems, comprising isolating a strain from  
XX microorganisms ecosystem, which is a dominant strain within the  
XX ecosystem, introducing a useful gene into the isolated strain and then  
XX returning the strain to the microorganisms ecosystem. The method is  
XX useful for fixing a useful gene e.g., phenyl hydroxylase gene to a strain  
XX dominant within a ecosystem, such as Comamonas sp. IN7 strain and its  
XX closely related strains. It is also useful for increasing phenol  
XX decomposition activity in a sludge, and for purifying environment  
XX contaminated with phenol. The current sequence represents a phenyl  
XX hydroxylase gene fragment.  
XX  
SQ Sequence 1463 BP; 374 A; 344 C; 465 G; 280 T; 0 U; 0 Other;  
Query Match 77.3%; Score 1126.4; DB 10; Length 1463;  
Best Local Similarity 87.2%; Pred. No. 0;  
Matches 1272; Conservative 0; Mismatches 176; Indels 10; Gaps 3;  
QY 3 TGAACGTGGCGCGCATCTTTACATGCAAGTGAACGGCAGACCGATGCTTCATCT 62  
DB 1 TGAACGTGGCGCGCATCTTTACATGCAAGTGAACGGCAGATG--GCTTCGCGCT 58  
QY 63 GGTGGGAGTGGGGGAGCGGTGATGATGATGCGAAGCTATCCAGAGAGGGGGGTAC 122  
DB 59 GATGGCGAGTGGCGAAGCGGTGATGATGATGCGAAGCTATGCGGTGAGGGGATAC 118  
QY 123 GCATCGAAGATGTGCTAATATCCGATATCTTAAGAGAGAAAGCAGGGATCGAAGA 182  
DB 119 TACTCGAAGATGATGCTAATATCCGATATCTAAGAGAGAAAGCAGGGATCGAAGA 178  
QY 183 CTTTGGCTTTTGGAGCGCGCATGCTGATTAAGTGTGGGTAAAGCCTTACCA 242  
DB 179 CTTGGCGCTACAGAGCGCGCTGATGCGATTAAGTGTGGGTAAAGCCTTACCA 238  
QY 243 AGCGGCGATCATGATGCTGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 302  
DB 239 AGCGGCGATCATGATGCTGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 298  
QY 303 CCAGACTCTTACAG 362  
DB 299 CCAGACTCTTACAG 358  
QY 363 GCAATGCCGCTGATGAG 422  
DB 359 GCAATGCCGCTGATGAG 418  
QY 423 AGGTTACGTAATTAATCGTACACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482  
DB 419 AGGCTCTTCTTAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478  
QY 483 CGTGCAG 542  
DB 479 CGTGCAG 538  
QY 543 AGGTTGGCAG 602  
DB 539 AGGTTGGCAG 598

QY 603 CGTTGAAACTACAAAGCTAGAGTGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662  
DB 599 CTTTGTGACTGGAAGCTGTGAGTGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658  
QY 663 AAATGCGTAGATATGAG 722  
DB 659 AAATGCGTAGATATGAG 718  
QY 723 CGCTATGACAGAAACGAG 782  
DB 719 CGCTATGACAGAAACGAG 778  
QY 783 AAACGATGCTCACTAGTGTGGCTTATGAGCTTGTGTAACGAGCTTAACGCTGAGAG 842  
DB 779 AAACGATGCTCACTAGTGTGGCTTATGAGCTTGTGTAACGAGCTTAACGCTGAGAG 838  
QY 843 TTGACCGCTGGGAGAGTACGCTGCAAGATTAAACTCAAGAGATTGACGAGAGACCGC 902  
DB 839 TTGACCGCTGGGAGAGTACGCTGCAAGATTGAACTCAAGAGATTGACGAGAGACCGC 898  
QY 903 ACAAGCGGTGATTAATGATGATTAATTCATGCAACGAGAGAGAGAGAGAGAGAGAG 962  
DB 899 ACAAGCGGTGATGATGATGATTAATTCATGCAACGAGAGAGAGAGAGAGAGAGAG 958  
QY 963 CATGTAGCGAATTTTCTAAGAGATGATGCT--TCGGGAAACGCTTAACAGAGTGTCT 1019  
DB 959 CATGTAGCGAATTTTCTAAGAGATGATGCT--TCGGGAAACGCTTAACAGAGTGTCT 1018  
QY 1020 GCATGCTGTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079  
DB 1019 GCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078  
QY 1080 CTTTGCATTAATTTGCACTATTTGGTGGGCACTTTAATGAGCTGCGGAGCAAAAC 1139  
DB 1079 CTTTGCATTAATTTGCTTACGAAA----GAGCACTTAATTTGAGCTGCGGAGCAAAAC 1133  
QY 1140 GGAGAGAGTGGGAGATGAGTCAAGTCTCATGAGCCCTTAATGAGGTAGGCTTACACGTA 1199  
DB 1134 GGAGAGAGTGGGAGATGAGTCAAGTCTCATGAGCCCTTAATGAGGTAGGCTTACACGTA 1193  
QY 1200 ATACAAATGCGCGTACAGAGAGGTTGCAACCGCGAGAGAGAGAGAGAGAGAGAG 1259  
DB 1194 ATACAAATGCGCGTACAGAGAGGTTGCAACCGCGAGAGAGAGAGAGAGAGAGAG 1253  
QY 1260 GTCTGATGCTGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCG 1319  
DB 1254 GTCTGATGCTGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCG 1313  
QY 1320 CGGATCAGATGTGCGGTGAATATCGTCCCGGCTTGTATACACACCGCCGTACACCA 1379  
DB 1314 TGGATCAGATGTGCGGTGAATATCGTCCCGGCTTGTATACACACCGCCGTACACCA 1373  
QY 1380 TGGAGTGGGTTTACACAG 1439  
DB 1374 TGGAGTGGGTTTACACAG 1433  
QY 1440 AATTTATGATCGGGGTG 1457  
DB 1434 GGGTTGTGATCGGGGTG 1451  
RESULT 39  
ABZ69300  
ID ABZ69300 standard; DNA; 1400 BP.  
XX  
XX AC ABZ69300;  
XX  
XX  
DT 11-AUG-2003 (first entry)  
XX  
XX J lividum 16S ribosomal RNA gene fragment #4.  
XX  
XX Osteoarthritis; antibacterial; Janthinobacterium; 16S RNA; gene;  
KM Osteopathic; antiarthritic; analgesic; anti-inflammatory; ds.

XX	Janthinobacterium lividum.
OS	
XX	WO2002102384-A1.
PN	
XX	27-DEC-2002.
PD	
XX	17-JUN-2002; 2002MO-GB002771.
XX	
PP	15-JUN-2001; 2001GB-00014672.
PR	
XX	(ORTH-) ORTHOGENICS AS.
XX	(GARD/) GARDNER R.
PA	
PI	El-Gewely MR;
XX	
DR	WPI, 2003-175199/17.
XX	
PT	Use of an antibacterial agent in the manufacture of a medicament for treating osteoarthritis.
PS	Example 1; Page 52-54; 89pp; English.
CC	The present invention relates to the use of an antibacterial agent in the production of a treatment for osteoarthritis. The bacteria causing osteoarthritis is Janthinobacterium lividum. The present sequence is a fragment of the 16S RNA coding sequence from J. lividum shown in the exemplification of the invention
CC	
XX	Sequence 1400 BP; 370 A; 310 C; 433 G; 287 T; 0 U; 0 Other;
SQ	
	Query Match            77.2%; Score 1125; DB 8; Length 1400;
	Best Local Similarity 89.0%; Pred. No. 0;
	Matches 1252; Conservative 0; Mismatches 145; Indels 10; Gaps 3.
Oy	27 CATGCAAGTGCAGACGGCAGACGCGATGCTTGTCATCTGTGTGCGAGTGGCGGACGGGTAG 86 
Dd	1 CATGCAAGTGCAGACGGCAGACGCGA--GCTTGCTGTGTGCGAGTGGCGGAACGGGTAG 58 
Oy	87 TAATGATCGGAACGTTATCAGAAGAGGGGGGTAACGATCGAAAGATGTGCTAATTACC 146 
Dd	59 TAATATATCGGAACGTTACTTAGAGTGGGGATTAAGTACGAAAGTTACGTTAATACC 118 
Oy	147 CATATACCTTAAGAGAGAAAGCAGGGGATCGAAGAAGCCTTGCGCTTTGGAAGCGGCCAT 206 
Dd	119 CATACATCTTAAGAGTGAAGTGGGGATCGAAGAAGCCTCATGCTGTGAGCGGCCAT 178 
Oy	207 GTCTGATTAAGTGTGTGTGGGTTAAAGCCTTACCAAGCGACATCACTAGTTGTTG 266 
Dd	179 ATCTGATTAGCTAGTTGTGTGGTTAAAGCCTTACCAAGCGATCGATCACTAGTTGTTG 238 
Oy	267 AAGAGACGACCAACCCACACTGGGACCTGAGCACAGGCCAGATCTTACCGGAGGACGACAG 326 
Dd	239 AGAGAGACCAACCCACACTGGAACCTGAGCACAGGTCAGATCTTACCGGAGGACGACAG 298 
Oy	327 TGAGGAATTTGGACATGGGCGCAGACCTGATCGACATGCCCCGTGAGTGAAGAAG 386 
Dd	299 TGGGGAATTTGGACATGGGCGAAGACCTGATCGACATGCCCCGTGAGTGAAGAAG 358 
Oy	387 CCTTGGGTTGTAAAGCTCTTTCACTGTCGAGAGAAAAGTTACGGTAAATAATCCGAC 446 
Dd	359 CCTTGGGTTGTAAAGCTCTTTGTCCAGGAGAAAACGGTGAAGAGCTTAATATCTTGT 418 
Oy	447 CATGAGCGTATCGACAGAGAGCAACGGCTTAATACGTCCAGACGCGCGTAAATAC 506 
Dd	419 AATGACGTTACTGAGAGATTAAGCACGGCTTAATACGTCCAGACGCGCGTAAATAC 478 
Oy	507 TAGGTCGAAGCGTTAATCGGAATTACTGGGCGTAAAGGGTGCAGAGCGGCTTTAAG 566 
Dd	479 TAGGTCGAAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCAGAGCGGTTTTTGAAG 538 
Oy	567 TCAGATGTGAATATCCCGGCTTAACCTGGGAATGTGCTTGAAGCTCAAGACTAGAG 626 

Db	539	TCTGATGTGAATATCCCCGGGCTCAACTGGGAATTCGATTGGAGACTGCGAAGCTAGTAAT	598
Qy	627	GTGCGACAGGAGAGGTGGAATTCATCTGTGTGACATGTAATTCGTAGATATGGAGAATC	686
Db	599	CTGGCAGAGGGGGGTAGAAATTCACGATGTGACAGTGAATGCGTAGATATGTGGAGAAC	658
Qy	687	ATCGATGCGGAAGCGACGCTCTCTGGGTTTAACTGACGCTCATATGCCAAGAAACGCTGGGG	746
Db	659	ACCGATGGCGAAGGACAGCCCCCTGGGTCAAGATTGACGCTCATGACGAAAGCGTGGGGA	718
Qy	747	GCAAAACAGGATTAAGTATCCCTGTATAGTCACAGCCCTAAACGATGTCAACTAGTTGTGGG	806
Db	719	GCAAAACAGGATTAAGTATCCCTGTATAGTCACAGCCCTTAACGATGTCTTACTAGTTGTCCG	778
Qy	807	CCTTATTAGCGTTGGTATCAAGAGCTAACGCGCTGAAAGTTGACCGGCTGGGAGTACGCTCG	866
Db	779	TCTTATTAGCTTGGTATCAAGAGCTAACGCGCTGAAAGTGAACGCTGGGAGTACGCTCG	838
Qy	867	CAAGATTTAAATCTCAAGAAATTGACGCGGACCCGACACAGCGGTGATTTATGTGATTA	926
Db	839	CAAGATTTAAATCTCAAGAAATTGACGCGGACCCGACACAGCGGTGATTTATGTGATTA	898
Qy	927	ATTGATGACGACGCGCAAAACCTTAACCTTACCTTGACATGTAAGGAAATTTCTAGAGATA	986
Db	899	ATTGATGACGACGCGCAAAACCTTAACCTTACCTTGACATGCGTGAATTCCTCGAGAGATT	958
Qy	987	GATTAGTGCT---TCGGGACGCTTACACAGGTGCTGATGAGCTGTCTGACGCTGTCTC	1043
Db	959	GGGGAGTGCTCGAAAGAACACAGTACACAGGTGCTGATGAGCTGTCTGACGCTGTCTC	1018
Qy	1044	GTGAGATGTTGGGTTAAGTCCCGCAACGAGGCAACCTTGTCAATTAATTCGATCATTT	1103
Db	1019	GTGAGATGTTGGGTTAAGTCCCGCAACGAGGCAACCTTGTCAATTAAGTTGC----TAC	1073
Qy	1104	GGTGGGCACTTTATGGAACCTGCGGTGACAAACCGGAGGAGGTGGGAGTACGTCGA	1163
Db	1074	GAAAGGGCACTTAATGAGACTGCGGTGACAAACCGGAGGAGGTGGGAGTACGTCGA	1133
Qy	1164	GTCCTCATGAGCCCTTATGAGGCTTACACGCTTATCAATGAGCGGTACAGAGGCTT	1223
Db	1134	GTCCTCATGAGCCCTTATGAGGCTTACACGCTTATCAATGAGTGTATCATACAGAGCGC	1193
Qy	1224	GCCAAACCGCAGGGGAGCTAATTCAGAAAGCGGCTGTATGTCGGATCGGAATTCGC	1283
Db	1194	GCCAAACCGCAGGGGAGCTAATTCAGAAAGGTATGTATGTCGGATGTGTATCTGTC	1253
Qy	1284	AATCGATCTCCGTGAAGTCGGAATGCGTAGTAATGCGGATCAGATGCGCGGTGAATA	1344
Db	1254	AATCGATCTCCGTGAAGTGTGAATGCGTAGTAATGCGGATCAGATGCGCGGTGAATA	1314
Qy	1344	CGTTCGCGGAGCTGTGTAACACCGCCCGTCAACCATGGAGATGGGTTTACACAGAGCA	1403
Db	1314	CGTTCGCGGAGCTGTGTAACACCGCCCGTCAACCATGGAGCGGGTTTACACAGAGTA	1373
Qy	1404	GGTAGCTTAACCGTAGAGAGGCGCTT	1430
Db	1374	GGTAGCTTAACCGTAGAGAGGCGCTT	1400
RESULT 40			
ADB61690			
ID	ADB61690 standard; DNA; 1488 BP.		
AC	ADB61690;		
XX	04-DEC-2003 (first entry)		
XX	16S rRNA of Burkholderia mallei DNA sequence.		
XX	enriching mRNA; high quality bacterial mRNA; bacterial gene expression;		
KM	poly-A tail; mRNA purification; oligo-dT capture;		
KM	prokaryote mRNA purification; bridging oligonucleotide; targeting region;		
KM	capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;		



Oy	CATGGAGTGGGTTTCACCAAGAAGGTACTTAACCGTAAGAGGGGCGTTCCACAGG	1437
Oy	1378 CATGGAGTGGGTTTCACCAAGAAGGTACTTAACCGTAAGAGGGGCGTTCCACAGG	1437
Dd	1376 CATGGAGTGGGTTTTACCAAGTGCGTTAETCTAACCCGAAGAGAGACGCTCACACAGG	1435
Oy	1438 TGAGATTCATGACTGGGGTG	1457
Dd	1436 TAGGATTATGACTGGGGTG	1455

## RESULT 41

ABZ69298  
ID ABZ69298 standard; DNA; 1400 BP.

AC ABZ69298;

DT 11-AUG-2003 (first entry)

DE J lividum 16s ribosomal RNA gene fragment #2.

**16S** **RNA**; **gene**; **Osteoarthritis**; **antibacterial**; **Janthinobacterium**; **KW**

[illegible][illegible]

XX  
FN

W0700Y7Z0ZM  
HY-60C70F4

21-DEC-2006  
PD  
XX

17-0000-2002; 2002MO-05002/11-  
PF  
XX

PR 15-JUN-2001; 2001GB-000148 / 2.  
XX

PA (ORTH-) ORTHOGENICS AS.  
PA (GARD/) GARDNER B.

XX  
PT E] -Gawo] v MB :

XX  
XX  
WPT: 2003-175199/17

XX  
XX  
Title of an anti-habour

PT treating osteoarthritis.

Example 1; Page 46-48; 89pp; English.

CC The present invention relates to the use of an antibacterial agent in the  
CC production of a treatment for osteoarthritis. The bacteria causing

osteoarthritis is *Janthinobacterium lividum*. The present sequence is a fragment of the 16S rRNA coding sequence from *J. lividum* shown in the

CC exemplification of the invention  
xx

SQ Sequence 1400 BP; 371 A; 307 C; 432 G; 290 T; 0 U; 0 Other;

Query Match	77.18;	Score 1123.4;	DB 8;	Length 1400;
Query Match	77.18;	Score 1123.4;	DB 8;	Length 1400;

Matches 1251; Conservative 0; Mismatches 146; Indels 10; Gaps 3

Qy 27 CATGCAAGTCGAACGGCAGACCGAATGCTTCATCTGGTGGCGAGTGGCGAACGGCTGAG 86  
Db 1 CATGTAAAGTCGAACGGCAGACCGA --GCTTGGCTCTGGTGGCGAAGTGGCGAAACGGGGTAG 58  
Qy 87 TAATGCATGGAACCGTATCCAGAAAGAGGGGGCGCTTAACGATCGAAABAATGCTAATACCG 146  
Db 59 TAATATATCGGAACGTAACCTTAGAGTGGGGGATTAACGTACGAAATTACGCTAATACCG 118  
Qy 147 CATATACCTTAAGAGGAAAGACAGGGGATCGAAGAAGCTTGCCTTTTGGAGCGGCCGAT 206  
Db 119 CATACGATCTTAAGGATGAAGTAGTGGGGGATGCAAGACCTTCATGCTGTGGAGGGCGGAT 178  
Qy 207 GTCGTATTAGCTAGTTGGTGGGGTAAAGGCTTACCAAGCGACGATCACTAGTTGGTCTG 266  
Db 179 ATCTGATTAGCTAGTTGGTGGGGTAAAGGCTTACCAAGGATGATCACTAGTGGTCTG 238

Qy	267	AGAGAGCAGCACGCCACACACTGAGGACTGAGACACGAGCCCAACATCTTACGGGAGGACGACG	326
Db	229	AGAGAGCAGCACGAGCCACACTGGAACCTGAGACACGAGTCCAACTCTACGGGAGCAGCAG	298
Qy	327	TGGGGAAATTTTGGACCAATGGGCGCAAGCTCGATCCAGCAATGCGCGCTGAGTGAAGAAG	386
Db	299	TGGGGAAATTTTGGACCAATGGGCGCAAGCTCGATCCAGCAATGCGCGCTGAGTGAAGAAG	358
Qy	387	CCCTTCGGGTTTAAAGCTCTTTGAGTCGAGAGAGAAAGTTACCGTAAATATCTGTACCC	446
Db	359	CCCTTCGGGTTTAAAGCTCTTTGAGTCGAGAGAGAAAGCGTGAAGACTAATATCTGTCT	418
Qy	447	CATGACGGTATCGACAGAGAGACCGCGCTAACTTACGTCCAGACCGCGGTAATACG	506
Db	419	AATGACGGTACTGAGAGATTAAGACCGGCTAACTTACGTCCAGACCGCGGTAATACG	478
Qy	507	TAGGGTGAAGCGTTAATCGGAATTACTGCGGCGTAAAGGGTGCAGCGCGCTTGTAG	566
Db	479	TAGGGTGAAGCGTTAATCGGAATTACTGCGGCGTAAAGCGTGCAGCGCGTTTTGTAG	538
Qy	567	TCAGATGTGAATCCCGGGCTTAACTGCGGAATTGCGTTGAAACTAACAGCTAAGT	626
Db	539	TCTGATGTGAATCCCGGGCTTAACTGCGGAATTGCGATTGGAACCTGCAAGCTTAAGT	598
Qy	627	GTGGCAGAGGGAGGTGGAATTCATGTGTACACATGAAATGCGTAGATATGGAAGAC	686
Db	599	CTGGCAGAGGGGGGTGGAATTCACGTGTACAGTGAATGCGTAGATATGGAAGAAC	658
Qy	687	ATCGATGCGAAGGAGCGCTCGTGGGTAACTGACGCTCATGACGAAAGCGTGGGGA	746
Db	659	ACCGATGCGAAGGAGCGCCCTCGTGGGTAAAGTGAACGCTCATGACGAAAGCGTGGGGA	718
Qy	747	GCAAAACAGGATTTAGATTAACCTGTAGTACAGCGCCCTAAACGATGCAACTAGTTGTGGG	806
Db	719	GCAAAACAGGATTTAGATTAACCTGTAGTACAGCGCCCTAAACGATGCTACTAGTTGTGGG	778
Qy	807	CCTTATTAAGCTTGGTAACTGAAGCTTAACGCTGGAAGTTGACCCGCTGGGGAGTACGGTGC	866
Db	779	TCTTATTAAGCTTGGTAACTGAAGCTTAACGCTGGAAGTGAAGCCGCTGGGGAGTACGGTGC	838
Qy	867	CAAGATTTAAATCTCAAGAGAAATTGACGGGGAGCCGCAACAACGGTGTATTAATGTGATTA	926
Db	839	CAAGATTTAAATCTCAAGAGAAATTGACGGGGAGCCGCAACAACGGTGTATTAATGTGATTA	898
Qy	927	ATTGATGCAACGCGAAAAACCTTACCTTACCTTGAACATGTAGCGAATTTCTAGAGATA	986
Db	899	ATTGATGCAACGCGAAAAACCTTACCTTACCTTGAACATGTAGCGAATTTCTTGAAGATC	958
Qy	987	GATTATGTCT---TGGGAGCGCTAAACAGGTGTGATATGGCTGTCTGACGTCTGTCT	1043
Db	959	AGGGAGTGTCTCAAAAGAGAACAGATACAGGTGTGATATGGCTGTCTGACGTCTGTCT	1018
Qy	1044	GTGAGATGTTGGTTTAAGTCCCGCAACGAGCGCAACCTTGTCAATTAATGCCATCATTT	1103
Db	1019	GTGAGATGTTGGTTTAAGTCCCGCAACGAGCGCAACCTTGTCAATTAATGTTC-----TAC	1073
Qy	1104	GCTTGGGCACTTTAAATGAGACTGCGGATGACAAACCGAGAGAGGTGGGATGACGTCAA	1163
Db	1074	GAAAGGCACTTAATGAGACTGCGGATGACAAACCGAGAGAGGTGGGATGACGTCAA	1133
Qy	1164	GTCTCTATGCGCTTAATGGGTATGGGCTTCAACCTTAATACATATGCGCGGTACAGAGGGTT	1222
Db	1134	GTCTCTATGCGCTTAATGGGTATGGGCTTCAACCTTAATACATATGAGAGCGCC	1199
Qy	1224	GCGAACCGCGAGAGGGGAGCTAATCTCAGAAAGCGCGTGTAGTCCGAGTCGGAATCTGC	1281
Db	1194	GCGAACCGCGAGAGGGGAGCTAATCTCAGAAAGGTATCTGTAGTCCGAGTTGTATCTGC	1251
Qy	1284	AACTCAGACTCGGTGAAGTTCGGAATCGCTAAGTAATGCGGATCAGCATGTCCGGGTGAATA	1343
Db	1254	AACTCAGACTCGGTGAAGTGTGAATGCTGTAGTAATGCGGATCAGCATGTCCGGGTGAATA	1311
Qy	1344	CGTTCCGGGCTTGTACACACCGCCGTACACCATATGGAGTGGGTTTACACAGAACCA	1403



DB 1314 CgTCCCGGCTCTTGTATACACCGCCCGCTACACCAATGGAGCGGTTTACCAAGAACTA 1373  
QY 1404 GGTAGTCTTAACCGTAAAGAGGCGCTT 1430  
DB 1374 GTAGCTTAACCGTAAAGAGGCGCTT 1400

RESULT 42  
ABV72366  
ID ABV72366 standard; DNA; 1452 BP.  
AC ABV72366;  
XX 16-DEC-2002 (first entry)  
DT  
XX  
DE Nucleotide sequence of a bacterial 16S rRNA gene fragment.  
XX  
KM 16S rRNA gene; beta proteobacteria; trichloroethylene; soil purification;  
KM underground water purification; phenol; ss.  
XX  
OS Unidentified.  
OS JP2002142756-A.  
XX  
PD 21-MAY-2002.  
XX  
PF 09-NOV-2000; 2000JP-00341513.  
XX  
PR 09-NOV-2000; 2000JP-00341513.  
XX  
XX (CHIK-) ZH CHIKU KANKYO SANGYO GIJITSU KENKYU.  
XX  
DR WPI; 2002-676987/73.  
XX  
PT Novel bacteria belonging to beta proteobacteria for purifying soil or  
PT underground water contaminated with phenol and trichloroethylene,  
PT decompose trichloroethylene in the presence of phenol under aerobic  
XX conditions.  
XX  
PS Claim 1; Page 5-6; 9pp; Japanese.  
XX  
CC The present sequence represents a fragment of the 16S rRNA gene of  
CC bacteria belonging to beta proteobacteria. The bacteria are capable of  
CC decomposing trichloroethylene with high efficiency. The bacteria are  
CC useful for purifying soil or underground water contaminated with phenol  
CC and trichloroethylene  
XX  
SQ Sequence 1452 BP; 375 A; 338 C; 448 G; 291 T; 0 U; 0 Other;

Query Match 77.0%; Score 1121.4; DB 6; Length 1452;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 1263; Conservative 0; Mismatches 176; Indels 6; Gaps 3;

QY 16 CATGCTTTACAGTGAAGTGAACGAGCAGCAGATGCTTGATCTGAGTGGAGTGGC 75  
DB 1 CATGCTTTACAGTGAAGTGAACGAGCAGCAGCAGATGCTTGATCTGAGTGGAGTGGC 58  
QY 76 GAGCGGCTGAGTATGATCGAAGCGATCCAGAAAGGCGGGTAAAGCATCGAAAGATG 135  
DB 59 GAAAGGCTGAGTATGATCGAAGCGATCCAGAAAGGCGGGTAAAGCATCGAAAGCTG 118  
QY 136 TGTCTAATACCGCATATCTCTTAAGAGAGAAAGCAGGGGATCGAAAGACTCTTGCTTTG 195  
DB 119 TGTCTAATACCGCATATGATCGAAGCGATCGAAAGCAGGGGATCGCAAGACTCTTGCGCAATG 178  
QY 196 GAGCGGCTGAGTATGATCGAAGCGATCCAGAAAGGCGGGTAAAGCGAAGCGAGATCG 255  
DB 179 GAGCGGCTGAGTATGATCGAAGCGATCGAAGCGGCTAAAGCGCTAAAGCGCTTCGATCTG 238  
QY 256 TAGTTGATCTGAGAGCAGCAGCAACTGGAGCTGAGACAGCGCCAGACTCTCTAGC 315  
DB 239 TAGCTGATCTGAGAGCAGCAGCAACTGGAGCTGAGACAGCGCCAGACTCTCTAGC 298

QY 316 GGAGCAGCAGTGGGGAAATTTTGAACAATGGGGCAAGCCCTGATCAGCAATGCCGCTG 375  
DB 299 GAGGCGCAGCAGTGGGGAAATTTTGAACAATGGGGCAAGCCCTGATCAGCAATGCCGCTG 358  
QY 376 AGTGAAGAAAGGCTTCGGGTTGTAAGCTCTTTCAGTTCGAGAAAGAAAGTTACGGTAA 435  
DB 359 CAGGATGAAGGCTTCGGGTTGTAAGCTCTTTCAGTTCGAGAAAGAAAGCTCTTTCAG 418  
QY 436 TAACTGTGACCAATGACGGTATCGACAGAAAGCAGCGGCTTAATCTAAGTCCAGAGCC 495  
DB 419 TACAGAAAGCTAATGACGGTATCGACAGAAAGCAGCGGCTTAATCTAAGTCCAGAGCC 478  
QY 496 GCGGTAATCTGAGGGTGAAGCGTTAAATCGAAATTAATCGGGGCTGTAAGGGTGGCAGGC 555  
DB 479 GCGGTAATCTGAGGGTGAAGCGTTAAATCGAAATTAATCGGGGCTGTAAGGGTGGCAGGC 538  
QY 556 GGCCTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTGGGAATTCGTTGAAACTAC 615  
DB 539 GGTATGTAAAGCAGTGTGAATCCCGGGCTTAACTGGGAATTCGTTGTAAGTCTG 598  
QY 616 AAAGCTAAGTGTGACAGAGGAGGAGTGAATTCATGTGTAGCAAGTGAATCGTGAAGA 675  
DB 599 ATAGCTAAGTGTGACAGAGGAGGAGTGAATTCATGTGTAGCAAGTGAATCGTGAAGA 658  
QY 676 TATGGAAGAACTGATGGCGAAGGCGAGCTCTGGGTTAACTGAGCGCTGATGACGA 735  
DB 659 TCGGAGGAACACCGATGGCGAAGGCGAGCTCTGGGTTAACTGAGCGCTGATGACGA 718  
QY 736 AAGCGTGGGAGCAGAAACAGAAATTAAGTAACTGATGATCCAGCGCTTAAACAGATGCA 795  
DB 719 AAGCGTGGGAGCAGAAACAGAAATTAAGTAACTGATGATCCAGCGCTTAAACAGATGCA 778  
QY 796 TAGTTGTTGGCCTTATTAAGCTTGTGTAACGAAGCTTAACGCTGAAGTTGACCGCTGG 855  
DB 779 TGGTGTGGGCTTCACTGACTGATGTAAGGCTTAACGCTGAAGTTGACCGCTGG 838  
QY 856 GAGTACGGTGGAGAAATTAAGTAACTGAAAGAAATTAAGCGGGGACCGGCAACAGCGGTGAT 915  
DB 839 GAGTACGGTGGAGAAATTAAGTAACTGAAAGAAATTAAGCGGGGACCGGCAACAGCGGTGAT 898  
QY 916 TATGTGATTAATTCGATGCAACGCGAAACCTTAACCTTACCTTACATGTAAGCAAT 975  
DB 899 GATGTGATTAATTCGATGCAACGCGAAACCTTAACCTTACCTTACATGTAAGCAAT 958  
QY 976 TTTCTAAGATTAATTAAGTCT--TCGGGAACGCTTAACAGAGTGTGATGCTGCT 1032  
DB 959 CCGCAGAGATGCTTAAGTGTGTAAGAAAGAAACCGTAACACAGGTGCTGATGCTGCT 1018  
QY 1033 CAGCTCGTGTGCGTGAAGTGTGTTAAGTCCCGCAAGAGCGCAACCTTGTCAATTAAT 1092  
DB 1019 CAGCTCGTGTGCGTGAAGTGTGTTAAGTCCCGCAAGAGCGCAACCTTGTCAATTAAT 1078  
QY 1093 TCCCATCATTTTGGTGGGCACTTTAATGAGACTGCGGCTGAACAAACCGGAGAAAGTGG 1152  
DB 1079 TCGTA-CATTCAAGTTGGGCACTTAATGAGACTGCGGCTGAACAAACCGGAGAAAGTGG 1137  
QY 1153 GATGAGTCAAGTCTCATGCGCTTATGAGGCTTACACGTAATTAATCAATGGCGCG 1212  
DB 1138 GATGAGTCAAGTCTCATGCGCTTATGAGGCTTACACGTAATTAATCAATGGCGCG 1197  
QY 1213 TACAGAGGTTGCCAAACCGGCGAGGGAGCTTAATCTCAGAAAGCGCGCTGATCCGGA 1272  
DB 1198 TACAAAGGTTGCCAAACCGGCGAGGGAGCTTAATCTCAGAAAGCGCGCTGATCCGGA 1257  
QY 1273 TCGGAGTCTGAACCTCGACTCGGTGAAGTGGAAATCGCTAATTCGCGATCAGCATGT 1332  
DB 1258 TCGGAGTCTGAACCTCGACTCGGTGAAGTGGAAATCGCTAATTCGCGATCAGCATGT 1317  
QY 1333 CCGGCTGAATACGTTCCCGGCTTGTACACACCGCCGCTCACACATGGAGTGGGTTT 1392  
DB 1318 CACGGTGAATACGTTCCCGGCTTGTACACACCGCCGCTCACACATGGAGTGGGTTT 1377







```
Db      1049 GTGCTGCATGCGTGTCTGACGCTGTGTCTGAGATGTTGGTTAACTCCCGCAACAGC 1108
Qy      1075 GCAACCCCTTGTCATTAAATGGCCATCATTTGGTGGGCACTTAAATGAGACTGCCGCTGAC 1134
Db      1109 GCAACCTTGTTCATTAGTTGCTCCATCATTCAGTTGGGCACTCTAATGAGACTGCCGCTGAC 1168
Qy      1135 AACCCGAGGAGAGGTGGGGATGACGTCAAGTCTCATGCGCTTAATGAGGCTTCAAC 1194
Db      1169 AAGCCGAGGAGAGGTGGGGATGACGTCAAGTCTCATGCGCTTAATGAGGCTTCAAC 1228
Qy      1195 ACGTAATCAATGCGCGCTACAGAGGGTGGCCAAACCGCGAGGGGAGCTAATCTCAAGAA 1254
Db      1229 ACGTCATACAAATGTCGGTATACAGAGGGTACCAAGCCGAGGCGGAGCCATCTCACAA 1288
Qy      1255 AGCGCGTGTAGTCCGATCCGATCGAGTCACTCGACTCCGTGAAGTCCGATCGTACT 1314
Db      1289 AACCATGTGATGTCGGATTCGACTCTGCAACTCGAGTCATGAAGTCCGATCGTACT 1348
Qy      1315 AATCGCGATCAGCATGTCTCGCGGTAAATACGTTCCCGGCTTTGTACACACCGCCCTCA 1374
Db      1349 AATCGCATGTGATGTCGGGTAAATACGTTCCCGGCTTTGTACACACCGCCCTCA 1408
Qy      1375 CACCATGGAGTGGGTTTACCAAGAGAGGTAGTCTTAACCTAAGAGGGGCGCTTGCA 1434
Db      1409 CACCATGGAGTGGGAGTACCAAGAGGTAGTAAACCAAGAGGTCCTTACCA 1468
Qy      1435 CGGTGATTCATGATCGGGGGT 1457
Db      1469 CGGTATGCTTATGATCGGGGTG 1491

RESULT 45
AAA81490_00
WP Sequence split into 15 fragments LOCUS AAA81490 Accession Aaa81490
WP Fragment Name Begin End
WP AAA81490_00 1 11000
WP AAA81490_01 100001 210000
WP AAA81490_02 200001 310000
WP AAA81490_03 300001 410000
WP AAA81490_04 400001 510000
WP AAA81490_05 500001 610000
WP AAA81490_06 600001 710000
WP AAA81490_07 700001 810000
WP AAA81490_08 800001 910000
WP AAA81490_09 900001 1010000
WP AAA81490_10 1000001 1110000
WP AAA81490_11 1100001 1210000
WP AAA81490_12 1200001 1310000
WP AAA81490_13 1300001 1410000
WP AAA81490_14 1400001 1437668
ID AAA81490 standard; DNA; 1437668 BP.
AC
XX AAA81490;
XX
XX 04-DEC-2000 (first entry)
XX
XX N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.
XX
XX N. meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
KM Meningococcus B; Memb; ds.
XX
OS Neisseria meningitidis.
XX
XX WO200022430-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US023573.
XX
XX 09-OCT-1998; 98US-0103794P.
PR 30-APR-1999; 99US-0132068P.
XX
```

```
PA (CHIR ) CHIRON CORP.
XX
XX Frazer CM, Hickey E, Peterson J, Tetrelin H, Venter JC;
PI Meisnani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;
PI Rappuoli R, Pizzi M;
XX
DR WPI: 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.
XX
XX Claim 7; Page 866-1272; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
CC Neisseria meningitidis Memb polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
CC
SQ Sequence 1437668 BP; 344338A; 353206C; 385074G; 355045T; 0U; 50cher;
Query Match 76.9%; Score 1119.8; DB 3; Length 110000;
Best Local Similarity 87.1%; Pred. NO. 0;
Matches 1275; Conservative 0; Mismatches 182; Indels 6; Gaps 4;
Qy 1 ATTGAAGCTGCGCGGATGCTTTACATGCAATCGACGCGAGCAC--GGATGCTTGC 58
Db 60992 ATTGAAGCTGCGCGGATGCTTTACATGCAATCGACGCGAGCACGAGGAAGCTTGC 61051
Qy 59 ATCT--GGTGGCGAGTGGCGGAGCGGTGATGATGATCGGAACGTATCCAGAGGGG 116
Db 61052 TTCTCGGCTGGCGAGTGGCGGAGCGGTGATGATGATCGGAACGTATCGAGTGGTGGG 61111
Qy 117 GGTAAAGCATGAAAGATGCTTAATACCGATATCTTAAGAGAGAAACAGGGGATC 176
Db 61112 GATTAAGTATGAAAGATCACTAATACCGATATCTTAAGAGAGAAAGAGGGAGACC 61171
Qy 177 GAAAGACTTGCGCTTTTGAAGCGGCGATGCTGATTAAGTATGCTTGGGTAAAGGC 236
Db 61172 TTGGGCGCTTGCGCTTATCGAGCGGCGCATATCTGATTAAGTATGCTTGGGTAAAGGC 61231
Qy 237 CTACCAAGGCGACATGATGATGTTGTTGAGAGGACGACGACCACTGGGACTGAGA 296
Db 61232 CTACCAAGGCGACATGATGATGTTGTTGAGAGGATGATCGGACCACTGGGACTGAGA 61291
Qy 297 CACGCGCCAGACTCTTACGCGGAGCGAGAGTGGGAAATTTTGAACATGGGCGCAAGCT 356
Db 61292 CACGCGCCAGACTCTTACGCGGAGCGAGAGTGGGAAATTTTGAACATGGGCGCAAGCT 61351
Qy 357 GATCCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGTAAGTCTTTCACTGACGAG 416
Db 61352 GATCCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGTAAGTCTTTGTCAGAGG 61411
```



QY 717 CACTAGCGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTGTGATGCCA 776  
| | | | |  
Db 108462 CACTAGCGCTCATGACCGAAAGCGTGGGATAGCAAAAGATTAGATACCTGTGATGCCA 108521  
| | | | |  
QY 777 CGCCCTTAAAGATGTCACTAGTGTGGGCTTATTA-GGCTTGGTAAAGAGCTAACG 835  
| | | | |  
Db 108522 CGCCCTTAAAGATGTCACTAGTGTGGGCAACCTGATGTGCTGTAGCTAGCTAACG 108581  
| | | | |  
QY 836 CGTGAATTTGACCGCTGGGGAGTACGGTGCAGATTTTAAACTCAAGAGATTTGACGGG 895  
| | | | |  
Db 108582 CGTGAATTTGACCGCTGGGGAGTACGGTGCAGATTTTAAACTCAAGAGATTTGACGGG 108641  
| | | | |  
QY 896 GACCCGCAAGACGGGTGATTTATGTGATTAATTGATGCAACGCGAAAACTTACCTTA 955  
| | | | |  
Db 108642 GACCCGCAAGACGGGTGATTTATGTGATTAATTGATGCAACGCGAAAACTTACCTG 108701  
| | | | |  
QY 956 CCTTGAACATGTACGAAATTTTCTAGAGATAGATTAGTG-CTTGGGAAAGCTTAACACAG 1014  
| | | | |  
Db 108702 GTCTTGAACATGTACGAAATCTCCGAGACGAGAGAGTCCCTTGGGAGCGGTAAACAG 108761  
| | | | |  
QY 1015 GTGCTGATAGGCTGTCCGACAGCTGTGCGTAGAGTGTGGGTTAAGTCCGCAACGAGC 1074  
| | | | |  
Db 108762 GTGCTGATAGGCTGTCCGACAGCTGTGCGTAGAGTGTGGGTTAAGTCCGCAACGAGC 108821  
| | | | |  
QY 1075 GCAACCCCTTGTCAATTATTTGCGCATATTGGTTGGGCACTTTAAATGAGACTGCCGGTAC 1134  
| | | | |  
Db 108822 GCAACCCCTTGTCAATTATTTGCGCATATTGGGCACTTTAAATGAGACTGCCGGTAC 108881  
| | | | |  
QY 1135 AAACCGAGAGAGTGGGGATGACGTCAAGTCTTCATGCCCCCTTAATGGGTAGGGCTTAC 1194  
| | | | |  
Db 108882 AAACCGAGAGAGTGGGGATGACGTCAAGTCTTCATGCCCCCTTAATGAGACTGCCGGCTTAC 108941  
| | | | |  
QY 1195 ACGTAAATCAATATGCGGTGACAGAGGGTTCGCAACCGCGAGGGGGAGCTAATCTCAGAA 1254  
| | | | |  
Db 108942 ACGTAAATCAATATGCGGTGACAGAGGGTTCGCAACCGCGAGGGGGAGCTAATCTCAGAA 109001  
| | | | |  
QY 1255 AGCGCGTGTAGTCCGATCCGATCCGAGTCTGCAACTCGACTCCGTGAAGTGGATCGTAGT 1314  
| | | | |  
Db 109002 AAGCGATGTAGTCCGATCCGATCCGAGTCTGCAACTCGACTCCGTGAAGTGGATCGTAGT 109061  
| | | | |  
QY 1315 AATGCGGATCAGCATGTGCGGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCTGCA 1374  
| | | | |  
Db 109062 AATGCGGATCAGCATGTGCGGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCTGCA 109121  
| | | | |  
QY 1375 CACCATGGGATGGGTTTCAACAGAAAGAGAGTCTAACGTTAAGAGAGGGGGCTTTCACA 1434  
| | | | |  
Db 109122 CACCATGGGATGGGTTTCAACAGAAAGAGTCTAACGTTAAGAGAGGGGGCTTTCACA 109181  
| | | | |  
QY 1435 CGGTGAGATTCATGACTGGGGTG 1457  
| | | | |  
Db 109182 CGGTGATCTTCAATGACTGGGGTG 109204  
| | | | |

## RESULT 47

AAAB1490.03

Continuation (4 of 15) of AAAB1490 from base 300001 (N. meningitidis B full length genom

Sequence split into 15 fragments LOCUS AAAB1490 Accession AaaB1490

WP	Fragment Name	Begin	End
WP	AAAB1490_00	1	110000
WP	AAAB1490_01	100001	210000
WP	AAAB1490_02	200001	310000
WP	AAAB1490_03	300001	410000
WP	AAAB1490_04	400001	510000
WP	AAAB1490_05	500001	610000
WP	AAAB1490_06	600001	710000
WP	AAAB1490_07	700001	810000
WP	AAAB1490_08	800001	910000
WP	AAAB1490_09	900001	1010000
WP	AAAB1490_10	1000001	1110000
WP	AAAB1490_11	1100001	1210000
WP	AAAB1490_12	1200001	1310000
WP	AAAB1490_13	1300001	1410000

WP	AAAB1490_14	1400001	1437668
QY	1 ATTTGAACGCTGGGCGCATGCTTTTACACATGCAAGTGAACGCGAGCAGC--GGATGCTTGC 58		
Db	7742 ATTTGAACGCTGGGCGCATGCTTTTACACATGCAAGTGAACGCGAGCAGCAGAGAAAGCTTGC 7801		
QY	59 ATCT--GGTGGCAGTGGCGGACGGGTGAGTAATGCATCGAAAGTATCCAGAAAGAGGGG 116		
Db	7802 TTCTCGGTGGCGAGTGGCGGACCGGTGAGTAATGCATCGAAAGTATCCAGAAAGTATGGGGG 7861		
QY	117 GGTAGCGCATGAAAGATGTGCTTAATCCGCAATACTCTTAAGAGAAAGACGGGGATC 176		
Db	7862 GATTAAGTATGAAAGATGATGCTTAATCCGCAATACTCTTGAAGAGAAAGACGGGGATC 7921		
QY	177 GAAAGACCTTGGCGCTTTGGAGGGCCGATGTCTGATTAAGTGTGGGTAAAGGC 236		
Db	7922 TTGGGCGCTTGGCGCTTTGGAGGGCCGATGTCTGATTAAGTGTGGGTAAAGGC 7981		
QY	237 CTACCAAGGCGAGATCAGTATGTGTCTGAGAGACGACACGACCACTGCGACTGAGA 296		
Db	7982 CTACCAAGGCGAGATCAGTATGTGTCTGAGAGAGATGATCCGCACTGCGACTGAGA 8041		
QY	297 CACGGCCAGACTCTCTACGGAAGCGACAGTGGGAAATTTTGGACAAATGGGCGCAAGCCT 356		
Db	8042 CACGGCCAGACTCTCTACGGAAGCGACAGTGGGAAATTTTGGCAATGGGCGCAAGCCT 8101		
QY	357 GATTCACCAATGCCCGCTGAGTGAAGAGGCTTGGGTTGTAAAGCTTTTCAGTGCAG 416		
Db	8102 GATTCACCAATGCCCGCTGAGTGAAGAGGCTTGGGTTGTAAAGACTTTTTCAGTGCAGG 8161		
QY	417 AAGAAAGGTTTACCGGTAAATTAATGTAACCATGACCGATATCGACAGAAAGAACCGGC 476		
Db	8162 AAGAAAGGCTGTTGCTTAATATCAAGCGGTATGATGACGATACCTGAAGAAATTAAGCACCGGC 8221		
QY	477 TAACTAGCTGCACAGACGCCGGTAAATCGTATGGGTGCAAGCTTAAATTCGAAATTAACCTG 536		
Db	8222 TAACTAGCTGCACAGACGCCGGTAAATCGTATGGGTGCAAGCTTAAATTCGAAATTAACCTG 8281		
QY	537 GCGTAAAGGTCGCAAGCGCCGCTTGAAGTCAGATGGAATATCCCGGGCTTAAACCTGG 596		
Db	8282 GCGTAAAGGTCGCAAGCGCCGCTTGAAGTCAGATGGAATATCCCGGGCTTAAACCTGG 8341		
QY	597 GAATTCGCTTGAATCAACAAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTA 656		
Db	8342 GAATTCGCTTGAATCAACAAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTA 8401		
QY	657 GCAGTGAATGCGTAGAGATATGGAAGAAATGATGAGCGGAGCGAGCTCTCTGGGTTAA 716		
Db	8402 GCAGTGAATGCGTAGAGATATGGAAGAAATGATGAGCGGAGCGAGCTCTCTGGGTTAA 8461		
QY	717 CACTAGCGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTGTGATGCCA 776		
Db	8462 CACTAGCGCTCATGACCGAAAGCGTGGGATAGCAAAAGATTAGATACCTGTGATGCCA 8521		
QY	777 CGCCCTTAAAGATGTCACTAGTGTGGGCTTATTA-GGCTTGGTAAAGAGCTAACG 835		
Db	8522 CGCCCTTAAAGATGTCACTAGTGTGGGCAACCTGATGTGCTGTAGCTAGCTAACG 8581		
QY	836 CGTGAATTTGACCGCTGGGGAGTACGGTGCAGATTTTAAACTCAAGAGATTTGACGGG 895		
Db	8582 CGTGAATTTGACCGCTGGGGAGTACGGTGCAGATTTTAAACTCAAGAGATTTGACGGG 8641		
QY	896 GACCCGCAAGACGGGTGATTTATGTGATTAATTGATGCAACGCGAAAACTTACCTTA 955		
Db	8642 GACCCGCAAGACGGGTGATTTATGTGATTAATTGATGCAACGCGAAAACTTACCTTAC 8701		
QY	956 CCTTGAACATGTACGAAATTTTCTAGAGATAGATTAGTG-CTTGGGAAAGCTTAACACAG 1014		
Db	8702 GTCTTGAACATGTACGAAATCTCCGAGACGAGAGAGTGCCTTGGGAGCGGTAAACAG 8761		

QY 1015 GTCTGTCATGCTGTCTGTCAGCTCTGTCTGTGAGATGTTGGGTTAAAGTCCCCGCAAGCAGC 1074  
DB GTCTGTCATGCTGTCTGTCAGCTCTGTCTGTGAGATGTTGGGTTAAAGTCCCCGCAAGCAGC 8821  
QY 1075 GCAACCTTGTCTTAATTATGTCATCATTTGTTGGGCACTTTATGAGACTGCGCGGTAC 1134  
DB GCAACCTTGTCTTAATTATGTCATCATTTGTTGGGCACTTTATGAGACTGCGCGGTAC 8881  
QY 1135 AAACCGGAGAGAGTGGGAGTGAAGTCAGTCCATGAGCCCTTATGAGGAGGCTTAC 1194  
DB AAACCGGAGAGAGTGGGAGTGAAGTCAGTCCATGAGCCCTTATGAGGAGGCTTAC 8941  
QY 1195 ACCTAATACATGCGCGGTACAGAGGTTGCGCAACCCGCGAGGGGAGCTAATCTCAGAA 1254  
DB ACCTAATACATGCGCGGTACAGAGGTTGCGCAACCCGCGAGGGGAGCTAATCTCAGAA 9001  
QY 1255 AGGCGGTGTAATGTCGATCGGAGTCTGCAACTGCACTCGTGAAGTCGGAATCGTAGT 1314  
DB AGGCGGTGTAATGTCGATCGGAGTCTGCAACTGCACTCGTGAAGTCGGAATCGTAGT 9061  
QY 1315 AATCGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACAACCGCGGTCA 1374  
DB AATCGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACAACCGCGGTCA 9121  
QY 1375 CACCATGGAGTGGGTTTACACAGAACAGTAGTCTAACCCTAAGAGAGGGCGCTTGCCA 1434  
DB CACCATGGAGTGGGTTTACACAGAACAGTAGTCTAACCCTAAGAGAGGGCGCTTGCCA 9181  
QY 1435 CGGTGAGATTCATGACTGGGGTG 1457  
DB CGGTGAGATTCATGACTGGGGTG 9204

## RESULT 48

AAAB1489\_2/c

Continuation (3 of 9) of AAAB1489 from base 200001 (N. meningitidis partial DNA sequence)

MP Sequence Eplc info 9 Fragments LOCUS AAAB1489 Accession AaAB1489

MP	Fragment Name	Begin	End
MP	AAAB1489_0	1	110000
MP	AAAB1489_1	100001	210000
MP	AAAB1489_2	200001	310000
MP	AAAB1489_3	300001	410000
MP	AAAB1489_4	400001	510000
MP	AAAB1489_5	500001	610000
MP	AAAB1489_6	600001	710000
MP	AAAB1489_7	700001	810000
MP	AAAB1489_8	800001	837096

Query Match 76.9%; Score 1119.8; DB 3; Length 110000;

Best Local Similarity 87.1%; Pred. No. 0; Mismatches 182; Indels 6; Gaps 4;

Matches 1275; Conservative 0;

QY 1 ATTTGAAGCTGGCGGAGCTTTTACATGCAAGTCGAAACGGCAGCAC--GGATGCTTGC 58  
DB ATTTGAAGCTGGCGGAGCTTTTACATGCAAGTCGAAACGGCAGCACAGAAAGCTTGC 95223  
QY 59 ATCT--GGTGGCAGTGGCGGACGGGTGATGATGATCGGAACGTATCCAGAAAGAGGG 116  
DB ATCT--GGTGGCAGTGGCGGACGGGTGATGATGATCGGAACGTATCCAGAAAGAGGG 95163  
QY 117 GGTACCCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATC 176  
DB GGTACCCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATC 95103  
QY 177 GAAAGACCTTGCGCTTTTGGAGCGGCGGATGTCATTAAGTATGTTGGGAGTAAAGC 236  
DB TTGCGGCTTGCGCTTTTGGAGCGGCGGATGTCATTAAGTATGTTGGGAGTAAAGC 95043  
QY 237 CTACCAAGGCGACATCAGTAGTTGATCTGAGAGAGACGACGACACTGGAGCTGAGA 296  
DB CTACCAAGGCGACATCAGTAGTTGATCTGAGAGAGATGATCCGACACTGGAGCTGAGA 94983

QY 297 CACGGCCGAGACTCTTAACGGAGAGCGACAGTGGGAAATTTTGGACAATGGGCGACAGCCT 356  
DB CACGGCCGAGACTCTTAACGGAGAGCGACAGTGGGAAATTTTGGACAATGGGCGACAGCCT 94923  
QY 357 GATTCAGCAATGCCGCTGATGTAAGAGGCTTGGGTTGTAAGCTTTTCAATCGAG 416  
DB GATTCAGCAATGCCGCTGATGTAAGAGGCTTGGGTTGTAAGCTTTTGTTCAGGG 94863  
QY 417 AAGAAAGGTTACGTTAATTCGTGACCCATGACCGTATGACAGAGAAAGCACCGGC 476  
DB AAGAAAGGCTGTTGTTAATTCAGCGGCTGATGACCGTATGACAGAGAAAGCACCGGC 94803  
QY 477 TAACTAGTCCAGACAGCCGCTAATACGTAAGGCTGCAAGGCTTATCGAATTAAC 536  
DB TAACTAGTCCAGACAGCCGCTAATACGTAAGGCTGCAAGGCTTATCGAATTAAC 94743  
QY 537 GCGTAAAGGCTGCGCAGCGGCTTGTAAAGTCAATGTGAATCCCGGGCTTAACCTGG 596  
DB GCGTAAAGGCTGCGCAGCGGCTTGTAAAGTCAATGTGAATCCCGGGCTTAACCTGG 94742  
QY 597 GAATTCGCTTTGAAACTAAGCTAAGATGTGGCAGAGGAGGTGGAATTCATGTTGTA 656  
DB GAATTCGCTTTGAAACTAAGCTAAGATGTGGCAGAGGAGGTGGAATTCATGTTGTA 94682  
QY 657 GCAGTAAATGCGTAGAGATATGAAAGAACATCGATGCGAAGGCGACGCTCTGGGTTAA 716  
DB GCAGTAAATGCGTAGAGATATGAAAGAACATCGATGCGAAGGCGACGCTCTGGGTTAA 94563  
QY 717 CACTGACGCTCATGACAGAAAGCGTGGGAGCAACAGATTAATCCTGTGATGCA 776  
DB CACTGACGCTCATGACAGAAAGCGTGGGAGCAACAGATTAATCCTGTGATGCA 94503  
QY 777 CGCCCTAAACGATGTCATGATGTTGGGCTTAA--GGCTGGTAAAGAACTTAAC 835  
DB CGCCCTAAACGATGTCATGATGTTGGGCTTAA--GGCTGGTAAAGAACTTAAC 94443  
QY 836 CGTGAAGTTGACCGGCTGGGAGTACGTCGCAAGATTAACTCAAGAAATTGAACGG 895  
DB CGTGAAGTTGACCGGCTGGGAGTACGTCGCAAGATTAACTCAAGAAATTGAACGG 94442  
QY 896 GACCCGACAAAGCGGTGATATGATGATTAATTCGATCAACCGCAAAACCTTAACCTA 955  
DB GACCCGACAAAGCGGTGATATGATGATTAATTCGATCAACCGCAAAACCTTAACCTA 94382  
QY 956 CCTTGAACATGACGCAATTTTCTAGATGATGTTGTC--CTTGGGAAACGCTAACACAG 1014  
DB CCTTGAACATGACGCAATTTTCTAGATGATGTTGTC--CTTGGGAAACGCTAACACAG 94322  
QY 1015 GTCTGTCATGCTGTGTCAGCTCTGTGTCGATGTTGGTTAAGTCCCGCAAGCAGC 1074  
DB GTCTGTCATGCTGTGTCAGCTCTGTGTCGATGTTGGTTAAGTCCCGCAAGCAGC 94203  
QY 1075 GCAACCTTGTCTTAATTGTCATCTTTGTTGGCACTTTAATGAGACTGCGGTGAC 1134  
DB GCAACCTTGTCTTAATTGTCATCTTTGTTGGCACTTTAATGAGACTGCGGTGAC 94202  
QY 1135 AAACCGGAGAGAGTGGGATGACGTCAGTCCATGAGCCCTTAAGGAGGCTTAC 1194  
DB AAACCGGAGAGAGTGGGATGACGTCAGTCCATGAGCCCTTAAGGAGGCTTAC 94083  
QY 1195 ACCTAATACATGAGCGCTGACAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAGAA 1254  
DB ACCTAATACATGAGCGCTGACAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAGAA 94023  
QY 1255 AGGCGGTGATGTCGGAATCGAGTGTGCAACTCGCTGGAATCGGAATCGCTAAGT 1314  
DB AGGCGGTGATGTCGGAATCGAGTGTGCAACTCGCTGGAATCGGAATCGCTAAGT 94022  
QY 1315 AATCGGGAATCAGCATGTGCGGCTGAAATACGTTCCCGGGCTTGTACACACGCGCGTCA 1374  
DB AATCGGGAATCAGCATGTGCGGCTGAAATACGTTCCCGGGCTTGTACACACGCGCGTCA 93962  
QY 1375 CACCATGGAGTGGGTTTACACGAAGCAGGTAGTCTAACCCTAAGAGAGGCGCTTGCCA 1434

Db 93902 CACCATGGAGTGGGGATACCAAGTAGTAGTAACCAAGAGATCGCTTACCA 93843  
Qy 1435 CGGTGATTCATGACTGGGGT 1457  
Db 93842 CGGTATGCTTCATGACTGGGGT 93820

## RESULT 49

AAAB1489\_6/c

Continuation (7 of 9) of AAAB1489 from base 600001 (N. meningitidis partial DNA sequence)

MP Sequence split into 9 fragments LOCUS AAAB1489 Accession AaaB1489

MP	Fragment Name	Begin	End
MP	AAAB1489_0	1	110000
MP	AAAB1489_1	100001	210000
MP	AAAB1489_2	200001	310000
MP	AAAB1489_3	300001	410000
MP	AAAB1489_4	400001	510000
MP	AAAB1489_5	500001	610000
MP	AAAB1489_6	600001	710000
MP	AAAB1489_7	700001	810000
MP	AAAB1489_8	800001	837096

Query Match 76.9%; Score 1119.8; DB 3; Length 110000;  
Best Local Similarity 87.1%; Pred. No. 0; Mismatches 182; Indels 6; Gaps 4;  
Matches 1275; Conservative 0;

Qy 1 ATGGAAGCTGGGGCGGATGCTTTACATGCAAGTCGACCGGACAC--GGATGCTTGC 58  
Db 102153 ATTGAAGCTGGGGCGGATGCTTTACATGCAAGTCGACCGGACACGAGAAAGCTTGC 102094  
Qy 59 ATCT--GGTGGCAGTGGCGGACCGGTGATGATGATCGGACGTATCCAGAAAGGGG 116  
Db 102093 TTCTCGGGTGGCGAGTGGCGAAGCGGTGATGATGATCGGACGTATCGAGTATGAGG 102034  
Qy 117 GGTAAAGCATGAAAGATGCTGTAATACCGCATATCTCTAAGAGAGAAACAGGGATC 176  
Db 102033 GATAATCATGATGAAAGATCAGTATATCCGATACGCTTGAAGAGAAACAGGGAGCC 101974  
Qy 177 GAAAGACCTTGGCGCTTTGAGCGGGCGGATGCTGATGATGATGATGATGATGATGATG 236  
Db 101973 TTCCGGCTTGGCGCTTATTCAGCGGGCGGATGCTGATGATGATGATGATGATGATG 101914  
Qy 237 CTACCAAGGCGGACATCAGTATGTTGCTGAGAGAGACGACCACTGGGACTGAGA 296  
Db 101913 CTACCAAGGCGGACATCAGTATGTTGCTGAGAGAGATGATCCGACACTGGGACTGAGA 101854  
Qy 297 CAGGCGGCGACTCTTACGAGGAGCAGAGTGGGAAATTTTGGACAAATGGCGCAAGCT 356  
Db 101853 CAGGCGGCGACTCTTACGAGGAGCAGAGTGGGAAATTTTGGACAAATGGCGCAAGCT 101794  
Qy 357 GATCAGCAATGCGCGGCGGATGAGTGAAGAGGCTTCGGGTTGAAAGCTCTTTCAGTCCAG 416  
Db 101793 GATCAGCAATGCGCGGCGGATGAGTGAAGAGGCTTCGGGTTGAAAGCTCTTTCAGTCCAG 101734  
Qy 417 AAGAAAGGTTACGGTAAATATATCTGTAACCATGACGGTATCGACAGAAAGCAACCGGC 476  
Db 101733 AAGAAAGGTTACGGTAAATATATCTGTAACCATGACGGTATCGTAAAGATTAAGCACCGGC 101674  
Qy 477 TAACTACGTCGACAGACCGCGGTTAATAGTGGTGCAGCGTTAATCGGAATTAACCTGG 536  
Db 101673 TAACTACGTCGACAGACCGCGGTTAATAGTGGTGCAGCGTTAATCGGAATTAACCTGG 101614  
Qy 537 GCGTAAAGGTCGCGAGCGGCGCTGTAAAGTCAATGTAATCCCGGGCTTAACTGG 596  
Db 101613 GCGTAAAGGTCGCGAGCGGCGCTGTAAAGTCAATGTAATCCCGGGCTTAACTGG 101554  
Qy 597 GAATTCGCTTTGAAACTAAGAGTGAAGTGTGGCAGAGAGGAGTGAATTCATGTGTA 656  
Db 101553 GAATTCGCTTTGAAACTAAGAGTGAAGTGTGGCAGAGAGGAGTGAATTCATGTGTA 101494  
Qy 657 GCAATGAAATGCTGATGATGATGAGAAATCATTCATGTCGAAAGGAGGCTCTCTGGGTTAA 716

Db 101493 GCAGTAAATGCTAGAGATGTGGAGAAATACCGATGGCGAAAGGACCTCTGGGACAA 101434  
Qy 717 CACTGACCTCATGACGAAAGGCTGGGAGCAACAGATTGATACCTGGTAGTCCA 776  
Db 101433 CACTGACCTCATGACGCGGAGGAGTGGAGCAACAGATTGATACCTGGTAGTCCA 101374  
Qy 777 CGCCCTTAAAGATGTAAGTATGTTGGGCTTTATTA--GGCTTGGTAAGAACTAACG 835  
Db 101373 CGCCCTTAAAGATGTAAGTATGTTGGGCTTTATTA--GGCTTGGTAAGAACTAACG 101314  
Qy 836 CGTGAAGTTGACCGCTGGGAGTACGCTGCAAGATTAACTCAAGAAATTTGACGG 895  
Db 101313 CGTGAAGTTGACCGCTGGGAGTACGCTGCAAGATTAACTCAAGAAATTTGACGG 101254  
Qy 896 GACCCGCAAGCGGCTGATTAATGATTAATTCATGCAAGCGGAAAACCTTACTTA 955  
Db 101253 GACCCGCAAGCGGCTGATTAATGATTAATTCATGCAAGCGGAAAACCTTACTTA 101194  
Qy 956 CCCTTGCATGTAGCGCAATTTTCTAGATAGATTAGTG--CTTCGGGAACGCTAACAG 1014  
Db 101193 CCCTTGCATGTAGCGCAATTTTCTAGATAGATTAGTG--CTTCGGGAACGCTAACAG 101134  
Qy 1015 GTGCTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074  
Db 101133 GTGCTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 101074  
Qy 1075 GCAACCTTGTCAATTAATGCTCATTTGTTGGGCACTTTAATGACCTGCCGTGAC 1134  
Db 101073 GCAACCTTGTCAATTAATGCTCATTTGTTGGGCACTTTAATGACCTGCCGTGAC 101014  
Qy 1135 AAACCGGAGAGGTCGGGAGTGAAGTCAATGCTCATGCGCTTATGATGATGATGATGATG 1194  
Db 101013 AAACCGGAGAGGTCGGGAGTGAAGTCAATGCTCATGCGCTTATGATGATGATGATG 100954  
Qy 1195 ACCTAATACATGCGCGCTACAGAGGCTTCCAAACCGCGAGGGGAGCTAATCTCAAA 1254  
Db 100953 ACCTAATACATGCGCGCTACAGAGGCTTCCAAACCGCGAGGGGAGCTAATCTCAAA 100894  
Qy 1255 AGCGGCTGATGCTCGGATGAGTGTGCAACTGCACTCCGTGAAGTCGGAATGCTGAT 1314  
Db 100893 AGCGGCTGATGCTCGGATGAGTGTGCAACTGCACTCCGTGAAGTCGGAATGCTGAT 100834  
Qy 1315 AATGCGGATCAGATGCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1374  
Db 100833 AATGCGGATCAGATGCTCGGATGATGATGATGATGATGATGATGATGATGATGATG 100774  
Qy 1375 CACCATGAGTGGGTTTCAACAGAACAGTACTTAACCGTAAAGAGGGCGCTTGCCA 1434  
Db 100773 CACCATGAGTGGGTTTCAACAGAACAGTACTTAACCGTAAAGAGGGCGCTTGCCA 100714  
Qy 1435 CGGTGATTCATGACTGGGGTG 1457  
Db 100713 CGGTGATTCATGACTGGGGTG 100691

## RESULT 50

AAAB1489\_7/c

Continuation (8 of 9) of AAAB1489 from base 700001 (N. meningitidis partial DNA sequence)

MP Sequence split into 9 fragments LOCUS AAAB1489 Accession AaaB1489

MP	Fragment Name	Begin	End
MP	AAAB1489_0	1	110000
MP	AAAB1489_1	100001	210000
MP	AAAB1489_2	200001	310000
MP	AAAB1489_3	300001	410000
MP	AAAB1489_4	400001	510000
MP	AAAB1489_5	500001	610000
MP	AAAB1489_6	600001	710000
MP	AAAB1489_7	700001	810000
MP	AAAB1489_8	800001	837096

Query Match 76.9%; Score 1119.8; DB 3; Length 110000;  
Best Local Similarity 87.1%; Pred. No. 0; Mismatches 182; Indels 6; Gaps 4;  
Matches 1275; Conservative 0;



```
QY 1 ATTTGAACGCTGGCGGATGCTTTTACACATGCAATGGAACGGGACAC--GGATGCTTGC 58
DB 2153 ATTTGAACGCTGGCGGATGCTTTTACACATGCAAGTCGGACGGGACGACAGAGAGGCTTGC 2094
QY 59 ATCT--GGTGGGAGATGGCGGAGCGGGTGAATATGCAATCGGAACGTAATCCAGAGAGGGG 116
DB 2093 TTCTCGGGTGGCGAGATGGCGAACGGGTGAATACATATCGGAACGTAATCCAGAGATGGGG 2034
QY 117 GGTACGCAATCGAAGATGTCTAATATACCGCATATATCTTAAAGAGAAAGCAGGGGATC 176
DB 2033 GATTAATGATCGAAGATGATGATATACCGCATATACGCTTGAAGAGAAAGCAGGGGACC 1974
QY 177 GAAAGACCTTGGCGCTTTTGGAGCGGCCGATGTCTGATTAAGTCTAGTTGGTGGGTTAAAGC 236
DB 1973 TTCTGGGCTTTGGCGCTATTTCCAGCGCGCATATGTGATTAAGTCTAGTTGGTGGGTTAAAGC 1914
QY 237 CTACCAAGGCGAGATCAGATGATTTGGTCTGAGAGAGACGACGACGACCACTGGGACTGAGA 296
DB 1913 CTACCAAGGCGAGATCAGATGATTTGGTCTGAGAGAGATGATCCGCACTGGGACTGAGA 1854
QY 297 CACGGCCACAGACTCTTACGGGAGGACGACAGTGGGGAATTTTGGACAATGGGCGCAAGCCT 356
DB 1853 CACGGCCACAGACTCTTACGGGAGGACGACAGTGGGGAATTTTGGACAATGGGCGCAAGCCT 1794
QY 357 GATCCAGCAATGCGCGGTGATGAGAGAGGCTTTCCGGTGTAAAGCTCTTTCAATCGAG 416
DB 1793 GATCCAGCAATGCGCGGTGATGAGAGAGGCTTTCCGGTGTAAAGGCTCTTTGTCAAGG 1734
QY 417 AAGAAAGGTTACGGTAAATATATGTAACCATGACCGGTATCGACAGAAAGACACCGGC 476
DB 1733 AAGAAAGGCTGTGTTGCTAATATACGCGGCTGATGACGCTACCTGMAAATATACCGGC 1674
QY 477 TAACTACGTGCGACAGACCGCGGTATATCGTAGGGTCAACGTTAATCGAAATTAATCG 536
DB 1673 TAACTACGTGCGACAGACCGCGGTATATCGTAGGGTCAACGTTAATCGAAATTAATCG 1614
QY 537 GCGTAAAGGCTGGCGAGCGCGCTTGTAACTCAGATGTGAAATCCCGGGCTTAAACCTCG 596
DB 1613 GCGTAAAGGCTGGCGAGCGCGTGTAACTTAAACGAGATGTGAAATCCCGGGCTTAAACCTCG 1554
QY 597 GAATGGGCTTGAACACTACAAAGCTAGATGTGGCAGAGGAGGTGGAATTTCCATGTGTA 656
DB 1553 GAATGGGCTTGTGAACACTGAGGTGACTCGAGTGTGTGAGAGGAGGTGGAATTTCCATGTGTA 1494
QY 657 GCAGTGAATGCGTAGAGATATGGAAGACATCGATGGCGAAGGCGACCTCTGGGTTAA 716
DB 1493 GCAGTGAATGCGTAGAGATGTGGAAGATATCGATGGCGAAGGCGACCTCTGGGACAA 1434
QY 717 CACTGACGCTGATGACGAAAGCGTGGGAGCAAAACGATTAAGTATCCCTGGTATGCCA 776
DB 1433 CACTGACGCTTATGATGCCGAAAGCGTGGTGTGCAAAACGATTAAGTATCCCTGGTATGCCA 1374
QY 777 CGCGCTAAAGATGTCAACTAGTTGTTGGGCTTATTA--GGCTTGTGTAAGCAAGCTTAAC 835
DB 1373 CGCGCTAAAGATGTCAACTAGTTGTTGGCAACTATTTGCTTGGTAGGTAGCTTAACG 1314
QY 836 CGTGAAGTTGACCGCTGGGAGATACGCTGCAAGATTAATACTCAAGGAATTTGACGGG 895
DB 1313 CGTGAAGTTGACCGCTGGGAGATACGCTGCAAGATTAATACTCAAGGAATTTGACGGG 1254
QY 896 GACCCGCAAAAGGGGTGATTAATGTGATTAATTTGATGCAACCGGAAAACTTACTTA 955
DB 1253 GACCCGCAAAAGGGGTGATTAATGTGATTAATTTGATGCAACCGGAAAACTTACTTA 1194
QY 956 CCTTGACATGATGCGAATTTTCTAGAGATAGATTAATG--CTTGCGGAGCGCTTAACAG 1014
DB 1193 GCTTGACATGATGCGAATCCTCGGAGACGAGAGAGTCTTGCGGAGCGCTTAACAG 1134
QY 1015 GTGCTGATGGCTGTCTGACAGTCTGTGCTGATGAGATGTGGTTAAGTCCCGCAACGAGC 1074
DB 1133 GTGCTGATGGCTGTCTGACAGTCTGTGCTGATGAGATGTGGTTAAGTCCCGCAACGAGC 1074
```

```
QY 1075 GCAACCTTGTGATTAATTCATATTCATATGTTGGGCACTTTAATGAGACTGCCGGTAC 1134
DB 1073 GCAACCTTGTGATTAATTCATATTCATATGTTGGGCACTTTAATGAGACTGCCGGTAC 1014
QY 1135 AAACCGAGAAAGTGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
DB 1013 AAACCGAGAAAGTGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
QY 1195 AGTAAATTAATGAGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1254
DB 953 AGTAAATTAATGAGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 894
QY 1255 AGCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
DB 893 AACCATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
QY 1315 AATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1374
DB 833 AATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
QY 1375 CACCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
DB 773 CACCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 714
QY 1435 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
DB 713 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 691
```

```
RESULT 51
AAF21613/c
ID AAF21613 standard; DNA; 172325 BP.
XX
AC AAF21613;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:114.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KM diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW de.
XX
OS Neisseria meningitidis.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Pizsa M, Hickey E, Peterson J, Tettelin H, Venter JC,
XX Maignani V, Galeotti C, Mora M, Ratti G, Scarfelli M, Scarlato V,
XX Rappoli R, Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
XX frames are used to detect, treat and prevent Neisserial infections.
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of Neisseria
XX meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
XX represent fragments of the NMB genomic sequence, as the sequence was too
XX long to go in a record on its own it was split into 8 sequences which
```





XX 30-APR-1999; 99US-0132068P.  
PR 08-OCT-1999; 99WO-US023573.  
PR 28-FEB-2000; 2000GB-00004695.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC,  
PI Meisigant V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,  
PI Rapinoli R, Frazer CM, Grandi G,  
XX WPI; 2000-647603/62.  
DR  
XX Neisseria meningitidis B full length genome sequence and open reading  
PT frames are used to detect, treat and prevent Neisserial infections.  
PS  
XX Claim 7; Appendix A; 692p; English.  
XX The present invention describes the full length genome of Neisseria  
CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613  
CC represent fragments of the NMB genomic sequence, as the sequence was too  
CC long to go in a record on its own it was split into 8 sequences which  
CC overlap each other at the beginning and end of each sequence by 49980 bp  
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of  
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of  
CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins  
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR  
CC primers which are used in the exemplification of the present invention.  
CC The NMB genome and fragments from it have antibacterial activity, and can  
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins  
CC and/or antibodies which binds to the proteins can be used in compositions  
CC for treating or preventing infection due to Neisserial bacteria or as a  
CC diagnostic reagent for detecting the presence of Neisserial bacteria or  
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,  
CC computer storage medium or computer databases can be used in a search to  
CC identify open reading frames (ORFs) or coding sequences within the NMB  
CC genome. The DNA sequences provide further opportunities to find antigenic  
CC or immunogenic proteins which are more effective in vaccines than the  
CC outer membrane proteins currently used  
XX  
SQ Sequence 349980 BP; 84410 A; 84863 C; 94187 G; 86520 T; 0 U; 0 Other;  
Query Match 76.9%; Score 1119.8; DB 3; Length 349980;  
Best Local Similarity 87.1%; Pred. No. 0;  
Matches 1275; Conservative 0; Mismatches 182; Indels 6; Gaps 4;  
QY 1 ATGGAAGCGTGGCGGATGCTTTACATGCAATGCAAGCGGAGCAGAC--GATGCTTGC 58  
DB 7742 ATTGAAGCGTGGCGGATGCTTTACATGCAATGCAAGCGGAGCAGACGATGCTTGC 7801  
QY 59 ATCT--GGTGGCGAGTGGCGGAGCGGTGAGTATGATCGAAGCTATCCAGAAAGAGGGG 116  
DB 7802 TTCTCGGATGGCGAGTGGCGGAGCGGTGAGTATGATCGAAGCTATCCAGAAAGCTGAGTGGG 7861  
QY 117 GGTAAAGCATGGAAGATGCTTAATACCGCATATCTCTAAGAGAGAAAGAGGGATC 176  
DB 7862 GATTAATCATGATCGAAGATGATACCGCATATGCTTGAAGAGAAAGAGAGGAGACC 7921  
QY 177 GAAAGACCTTGGCGCTTTGAGCGGCGGATGCTGATTAAGTATGTTGGGAGTAAAGGC 236  
DB 7922 TTGGGCGCTTGGCGCTTATTCAGCGGCGGATGCTGATTAAGTATGTTGGGAGTAAAGGC 7981  
QY 237 CTACCAAGGCGAGCATGATGTTGTCTGAGAGAGCAGCAGCCACACTGAGAGCTAGA 296  
DB 7982 CTACCAAGGCGAGCATGATGTTGTCTGAGAGAGCAGCAGCCACACTGAGAGCTAGA 8041  
QY 297 CAGGCGCCAGACTCTTACGGAAGCAGCAGTGGGAAATTTTGGCAATGGCGGAAACCT 356  
DB 8042 CAGGCGCCAGACTCTTACGGAAGCAGCAGTGGGAAATTTTGGCAATGGCGGAAACCT 8101  
QY 357 GATCCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGTAAGCTCTTTCAGTCAAG 416  
|||||

DB 8102 GATCCAGCCATGCGCGGTGCTGAAGAAAGCCTTGGGTTGTAAAGACTTTTGTCAAGG 8161  
QY 417 AAGAAAAGGTACGTAATTAATTCGTGACCCATGAGCGTATGACAGAAAGACACGGC 476  
DB 8162 AAGAAAAGGTGTTGTGTAATTCAGGGCTGATGACGGTATCTGAAGAAATGACACGGC 8221  
QY 477 TAACTACGTGCGCAGCAGCGCGGTAAATAGTAAAGGTGCAAGCGTTAATCGAATTAACG 536  
DB 8222 TAACTACGTGCGCAGCAGCGCGGTAAATAGTAAAGGTGCAAGCGTTAATCGAATTAACG 8281  
QY 537 GCGTAAAGGCGTGGCAGCGCGCTTGAAGTCAATGATGTAATCCCGGCTTAACCTG 596  
DB 8282 GCGTAAAGGCGTGGCAGCGCGCTTGAAGTCAATGATGTAATCCCGGCTTAACCTG 8341  
QY 597 GAATTCGCTTTGAATTAACCAAGCTAGTGTGCGAGAGGAGTGAATTCATGCTGTA 656  
DB 8342 GAATTCGCTTTGAATTAACCAAGCTAGTGTGCGAGAGGAGTGAATTCATGCTGTA 8401  
QY 657 GCAGTAAATGCGTAGATATGAAAGAACATCGATGCGGAGGCGCTCTGAGGTTAA 716  
DB 8402 GCAGTAAATGCGTAGATATGAAAGAACATCGATGCGGAGGCGCTCTGAGGTTAA 8461  
QY 717 CACTGACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTGATACCTGTGATGCCA 776  
DB 8462 CACTGACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTGATACCTGTGATGCCA 8521  
QY 777 CGCCCTAAACGATGTAATGATGTTGGGCTTATTA--GGCTTGTAAAGAACTTAACG 835  
DB 8522 CGCCCTAAACGATGTAATGATGTTGGGCTTATTA--GGCTTGTAAAGAACTTAACG 8581  
QY 836 CGTGAAGTTGACCGCGCTGGGAGTACGCTGCGAAGATTAACTCAAGAAATTGAACGG 895  
DB 8582 CGTGAAGTTGACCGCGCTGGGAGTACGCTGCGAAGATTAACTCAAGAAATTGAACGG 8641  
QY 896 GACCCGCAACAGCGGTGATTTATGATGATTAATTCATGCAACGCGAAACCTTACCTA 955  
DB 8642 GACCCGCAACAGCGGTGATTTATGATGATTAATTCATGCAACGCGAAACCTTACCTA 8701  
QY 956 CCTTGAACATGTAAGCAATTTTCTAGATGATGATTTGTTG--CTTGGGAAACGCTTAACAG 1014  
DB 8702 GTCTTAACATGTAAGCAATTTTCTAGATGATGATTTGTTG--CTTGGGAAACGCTTAACAG 8761  
QY 1015 GTGCTGATGAGCTGTGTCAGCTGCTGTGAGATGTTGGTTAAGTCCCGACAGAGC 1074  
DB 8762 GTGCTGATGAGCTGTGTCAGCTGCTGTGAGATGTTGGTTAAGTCCCGACAGAGC 8821  
QY 1075 GCAACCTTGTCAATTAATGCTATCAATTTGTTGGCACTTTAATGAGACTGCGGTGAC 1134  
DB 8822 GCAACCTTGTCAATTAATGCTATCAATTTGTTGGCACTTTAATGAGACTGCGGTGAC 8881  
QY 1135 AATCCGGAAGAGTGGGAGTACGTCAGTCCATGAGCCCTTATGGGTAAGGAGCTTAC 1194  
DB 8882 AATCCGGAAGAGTGGGAGTACGTCAGTCCATGAGCCCTTATGGGTAAGGAGCTTAC 8941  
QY 1195 ACCTAATTAACATGAGCGCGTACAGAGGTTGCCAACCAGGAGGAGGAGCTTAATCTAGAA 1254  
DB 8942 ACCTAATTAACATGAGCGCGTACAGAGGTTGCCAACCAGGAGGAGGAGCTTAATCTAGAA 9001  
QY 1255 AGCGCGTGTGATGTCGAGTCCGATGCGAGTCTGCACTCCGTAAGTCCGAATGCTGAT 1314  
DB 9002 AACCGGATCGTATGTCGAGTCCGATGCGAGTCTGCACTCCGTAAGTCCGAATGCTGAT 9061  
QY 1315 AATCGGAGATCAGCATGTCGCGGTGAATAGCTCCGCGGCTTGTACACACGCGCTCA 1374  
DB 9062 AATCGGAGATCAGCATGTCGCGGTGAATAGCTCCGCGGCTTGTACACACGCGCTCA 9121  
QY 1375 CACCATGAGAGTGGGTTTCCACAGAGCAGGTAGTCTAAACCGTAAAGAGGCGCTTGCCA 1434  
DB 9122 CACCATGAGAGTGGGTTTCCACAGAGCAGGTAGTAGTAAACCAACAGAGTCCGCTTACCA 9181  
QY 1435 CCGTGAAGATTGATGACTGGGGTG 1457  
DB 9182 CCGTGAAGATTGATGACTGGGGTG 9204  
|||||

RESULT 53  
AAFP21611/c  
ID AAFP21611 standard; DNA: 349980 BP.  
XX  
XX AAFP21611;  
XX  
XX 13-MAR-2001 (first entry)  
XX  
XX Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.  
XX  
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;  
KM diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
KM ds.  
XX  
XX Neisseria meningitidis.  
OS  
XX WO20006791-A1.  
PN  
XX 09-NOV-2000.  
PD  
XX 08-MAR-2000; 2000WO-US005928.  
PF  
XX 30-APR-1999; 99US-0132068P.  
PR 08-OCT-1999; 99WO-US023573.  
PR 28-FEB-2000; 2000GB-00004695.  
XX  
XX (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX Piazza M, Hickey E, Peterson J, Tettelin H, Venter JC,  
PI Maigiani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
PI Rappunli R, Frazer CM, Grandi G;  
XX WPI: 2000-647603/62.  
DR  
XX  
XX Neisseria meningitidis B full length genome sequence and open reading  
PT frames are used to detect, treat and prevent Neisserial infections.  
XX  
XX  
PS Claim 7; Appendix A; 692pp; English.  
XX  
XX The present invention describes the full length genome of Neisseria  
CC meningitidis B (NMB). The sequences in AAFP21544 and AAFP21607 to AAFP21613  
CC represent fragments of the NMB genomic sequence, as the sequence was too  
CC long to go in a record on its own it was split into 8 sequences which  
CC overlap each other at the beginning and end of each sequence by 49980 bp  
CC (i.e. the last 49980 bp of AAFP21544 is repeated at the beginning of  
CC AAFP21607, the last 49980 bp of AAFP21607 are repeated at the beginning of  
CC AAFP21608, and so on). AAFP21545 to AAFP21588 encode the Neisseria proteins  
CC given in AAB58550 to AAB58593, and AAFP21589 to AAFP21606 represent PCR  
CC primers which are used in the exemplification of the present invention.  
CC The NMB genome and fragments from it have antibacterial activity, and can  
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins  
CC and/or antibodies which binds to the proteins can be used in compositions  
CC for treating or preventing infection due to Neisserial bacteria or as a  
CC diagnostic reagent for detecting the presence of Neisserial bacteria or  
CC of antibodies raised to Neisserial bacteria. Computer, computer memory,  
CC computer storage medium or computer databases can be used in a search to  
CC identify open reading frames (ORFs) or coding sequences within the NMB  
CC genome. The DNA sequences provide further opportunities to find antigenic  
CC or immunogenic proteins which are more effective in vaccines than the  
CC outer membrane proteins currently used  
XX  
SQ Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 U; 0 Other;  
Query Match 76.9%; Score 1119.8; DB 3; Length 349980;  
Best Local Similarity 87.1%; Pred. No. 0;  
Matches 1275; Conservative 0; Mismatches 182; Indels 6; Gaps 4;  
QY 1 ATTTGAACGCTGGCGGCGATGCTTTACATGCAAGTGGCAAGCGGACAC--GGATGCTTGC 58  
DB 230511 ATTGAACGCTGGCGGCGATGCTTTACATGCAAGTGGCAAGCGGACACGAGAACTTGC 230452

QY 59 ATCT--GTTGCGAAGTGGCGGACCGGTGAGTATGATCGGAACGTATCCAGAGAGGG 116  
DB 230451 TTCTCGGCTGGCGGCGGCGAAGCGGTGAGTATGATCGGAACGTATCCAGAGAGTGGG 230392  
QY 117 GGTACGCAATCGAAGATGTGTATATACGCAATATCTCTAGAGAGAAAGCAGGGGATC 176  
DB 230391 GATTAAGTATCGAAGATATAGCTATATACGCAATACGCTTGAAGAGAAAGCAGGGGACC 230332  
QY 177 GAAAGACCTTGGCGCTTTGGAGCGGCCGATGTCTGATTAGCTAGTTGGTGGGCTAAAGGC 236  
DB 230331 TTGGGCGCTTGGCGCTTATTCGAGCGCGGATATCTGATTAGCTAGTTGGTGGGCTAAAGGC 230272  
QY 237 CTACCAAGGAGGAGATCAAGATGTGTGTGAGAGAGAGCAGACAGCACTGGGACTGAGA 296  
DB 230271 CTACCAAGGAGGAGATCAAGATGTGTGTGAGAGAGATATCCGCCACACTGGGACTGAGA 230212  
QY 297 CACGCGCCAGACTCTTACGAGGAGGAGCAGTGGGGAATTTTGACAAATGGGCGCAAGCCT 356  
DB 230211 CACGCGCCAGACTCTTACGAGGAGGAGCAGTGGGGAATTTTGACAAATGGGCGCAAGCCT 230152  
QY 357 GATTCAGCAATGCGCGCTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCACTCGAG 416  
DB 230151 GATTCAGCAATGCGCGCTGAGTGAAGAGGCTTCGGGTTGTAAGAGACTTTTGTCAAGG 230092  
QY 417 AAGAAAAGGTTACGCTTAATATCTGCAAGCCATGACGCTATCGACAGAGAGCAGCCGC 476  
DB 230091 AAGAAAAGGCTGTTGCTTAATATCTGCAAGCCATGAGCGCTATCGAAGAAATAGCACCGGC 230032  
QY 477 TAACTACGTCGACAGACCGCGGTAATACCTAGGGTGCAGACGTTAATCGGAATTAATCG 536  
DB 230031 TAACTACGTCGACAGACCGCGGTAATACCTAGGGTGCAGACGTTAATCGGAATTAATCG 229972  
QY 537 GCGTAAAGGCTGCGCAGCGCGCTTGTAGTCAAGATGTGAATCCCGGCTTAACCTGG 596  
DB 229971 GCGTAAAGGCTGCGCAGCGCGCTTGTAGTCAAGATGTGAATCCCGGCTTAACCGCG 229912  
QY 597 GAATTCGCTTGAACACTCAAGAGTATAGTGGCAGAGAGGAGGTGAATTCATGTGTA 656  
DB 229911 GAATTCGCTTGAACACTCAAGAGTATAGTGGTCAAGAGAGGAGGTGAATTCATGTGTA 229852  
QY 657 GCAAGTGAATGCGTAGAGATATGGAAGAACATGATGCGGACGCTCTCTGGGTTAA 716  
DB 229851 GCAAGTGAATGCGTAGAGATATGGAAGAACATGATGCGGACGCTCTCTGGGTTAA 229792  
QY 717 CACTGACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCCA 776  
DB 229791 CACTGACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCCA 229732  
QY 777 CGCCCTTAAGATGTCAATGATGTTGGGCTTATTA--GGCTTGTATGGAAGCTTAACG 835  
DB 229731 CGCCCTTAAGATGTCAATGATGTTGGGCTTATGATGCTTGTGTGATGCTTAACG 229672  
QY 836 CGTGAAGTGAACGCGCTGGGGAGTACGCTGCGCAAGATTTAAACTCAAGGAATTTAGCGG 895  
DB 229671 CGTGAAGTGAACGCGCTGGGGAGTACGCTGCGCAAGATTTAAACTCAAGGAATTTAGCGG 229612  
QY 896 GACCCGCAAGACCGGTGATTTATGATGATTTAATTCAGTCAACCGCAAAACCTTACTTA 955  
DB 229611 GACCCGCAAGACCGGTGATTTATGATGATTTAATTCAGTCAACCGCAAAACCTTACTTA 229552  
QY 956 CCTTGAATGATGCAATTTTCTAGAGATGATTAAGT--CTTGGGGAACCTTAAACAGAG 1014  
DB 229551 GCTTGAATGATGCAATTTTCTAGAGATGATTAAGT--CTTGGGGAACCTTAAACAGAG 229492  
QY 1015 GTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074  
DB 229491 GTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 229432  
QY 1075 GCAACCCCTTGTATTAATTTGCTATTTGGTGGGCACTTTAATGAGACTGCGGCTGAC 1134  
DB 229431 GCAACCCCTTGTATTAATTTGCTATTTGGTGGGCACTTTAATGAGACTGCGGCTGAC 229372

QY	1135	AAACCGGAGGAGCTGGGGATACGTCAGTCTCTATGGCCCTTATGGGTAGGGCTTCAC	1194
Db	229371	AAGCCGGAGGAAGTGGGGATGACGTCAGTCTCTATGGCCCTTATGACCAAGGCTTCAC	229312
QY	1195	ACGTATATCAATGCGCGCTPACAGAGGGTTGCCAACCCGAGAGGGAGAGCTAATCTCAGAA	1254
Db	229311	ACGTATATCAATGCTGTCGTCACAGAGGGTATGCCAAGCCGAGAGGGAGAGCTAATCTCAGAA	229252
QY	1255	AGCGCGTCTGTAATCCGGATCGAGATCTGCAACTGCACTCCGTGAACTCGGAATCGTAACT	1314
Db	229251	AACCGATCGTAACTCCGGATTTGCACTCTGCAACTCGAGTGCATGAAGTCGGAATCGTAACT	229192
QY	1315	AATCCGCGGATCAGCATGTGCGCGTGAATACGTTCCCGGCTCTGTACACACGCCGCCGTCA	1374
Db	229191	AATCCGCGGATCAGCATGTGCGCGTGAATACGTTCCCGGCTCTGTACACACGCCGCCGTCA	229132
QY	1375	CACCATGGGAGTGGGGTTTCCACGAGAGAGGTATGATTAACCGTAAAGGAGGAGCTTGGCCA	1434
Db	229131	CACCATGGGAGTGGGGGATTAACGAGAGTATGATTAACACAGAGAGTCCGCTTACCA	229072
QY	1435	CGGTGAGATTCATGACTGGGGGTG	1457
Db	229071	CGGTATGCTTCATGACTGGGGGTG	229049
RESULT 54			
AAFP21612/C			
ID	AAFP21612	standard; DNA; 349980 BP.	
XX	AAFP21612;		
AC	AAFP21612;		
XX			
DT	13-MAR-2001	(first entry)	
XX			
DE	Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.		
XX			
KW	Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;		
KM	diagnosis; antigen; detection; infection; gene therapy; antibacterial;		
KW	ds.		
OS	Neisseria meningitidis.		
XX			
PN	WO200066791-A1.		
XX			
PD	09-NOV-2000.		
XX			
PP	08-MAR-2000; 2000WO-US005928.		
XX			
PR	30-APR-1999; 99US-0132068P.		
XX			
PR	08-OCT-1999; 99WO-US023573.		
XX			
PR	28-FEB-2000; 2000GB-00004695.		
XX			
PA	(CHIR ) CHIRON CORP.		
XX			
PA	(GENO-) INST GENOMIC RES.		
XX			
P1	Pizza M, Hickey E, Peterson J, Tetelijn H, Venter JC,		
P1	Masignani V, Galeotti C, Mora M, Ratti G, Scarcella M, Scarlato V;		
P1	Rapnuoli R, Frazer CM, Grandi G;		
XX			
DR	WPI; 2000-647603/62.		
XX			
PT	Neisseria meningitidis B full length genome sequence and open reading		
XX	frames are used to detect, treat and prevent Neisserial infections.		
PS	Claim 7; Appendix A; 692pp; English.		
CC	The present invention describes the full length genome of Neisseria		
CC	meningitidis B (NMB). The sequences in AAFP21544 and AAFP21607 to AAFP21613		
CC	represent fragments of the NMB genomic sequence, as the sequence was too		
CC	long to go in a record on its own it was split into 8 sequences which		
CC	overlap each other at the beginning and end of each sequence by 49980 bp		
CC	(i.e. the last 49980 bp of AAFP21544 is repeated at the beginning of		
CC	AAFP21607, and the last 49980 bp of AAFP21607 are repeated at the beginning of		
CC	AAFP21608, and so on). AAFP21545 to AAFP21588 encode the Neisseria proteins		

	CC	given in AAB58550 to AAB58593, and AAP21589 to AAP21606 represent PCR
	CC	primers which are used in the exemplification of the present invention.
	CC	The NMB genome and fragments from it have antibacterial activity, and can
	CC	be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
	CC	and/or antibodies which binds to the proteins can be used in compositions
	CC	for treating or preventing infection due to Neisserial bacteria or as a
	CC	diagnostic reagent for detecting the presence of Neisserial bacteria or
	CC	of antibodies raised to Neisserial bacteria. Computer, computer memory,
	CC	computer storage medium or computer databases can be used in a search to
	CC	identify open reading frames (ORFs) or coding sequences within the NMB
	CC	genome. The DNA sequences provide further opportunities to find antigenic
	CC	or immunogenic proteins which are more effective in vaccines than the
	CC	outer membrane proteins currently used
SQ	XX	Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 U; 0 Other;
		Query Match 76.9%; Score 1119.8; DB 3; Length 349980;
		Best Local Similarity 87.1%; Pred. No. 0; Mismatches 182; Indels 6; Gaps 4;
		Matches 1275; Conservative 0;
OY	1	ATTGAAACCTTGCGGCATGCTTTTACACATGCAAGTCGAACGGCAGCAC--GGATGCTTGC 58
Dd	337382	ATTGAACCTTGCGGCATGCTTTTACATGCAAGTCGAACGGCAGCACAGAAGCTTGC 337323
OY	59	ATCT--GGTGGAGAGTGGCGGACGGGTGATGATCATCGAACCTATTCAGAAAGGGG 116
Dd	337332	TTCCTGGGTGGCGAGTGGCGGAACGGGTGATGATCATTCGGAACCTACGACTAGTGGGG 337263
OY	117	GGTAAAGCATGGAAGAAGTGTCTAATACCAGTAATACCTCTAAGAGGAAGCAAGCAGGGATC 176
Dd	337262	GATAACTGATGGAAGAAGTGTATATACCGCATACCTCTTGAGAGAGAAAGCAGGGAGCC 337203
OY	177	GAAGAAGCTTGGCGCTTTTGGAGCGGCCGATGCTGATTAGCTAGTTGGTGGGTAAAGGC 236
Dd	337202	TTCGGGCTTGGCGATATTCGAGCGGCCGATATCTGATTAGTAGTTGGTGGGTAAAGGC 337143
OY	237	CTACCAAGGCGACGATCTCACTAGTATGTTGCTTGAGAGGACGACACCACTCTGGGACTGANA 296
Dd	337142	CTACCAAGGCGACGATCTAGTAGCGGGTCTGAGAGGATGATTCGCACTCTGGGACTGANA 337083
OY	297	CACGGCCGAGACTCCTACGGGAGGAGAGAGTGGGGAAATTTTGGACAATGGGGCGAACGCT 356
Dd	337082	CACGGCCGAGACTCCTACGGGAGGAGAGAGTGGGGAAATTTTGGACAATGGGGCGAACGCT 337023
OY	357	GATCAGCAATGCCCGCTGAGTGAAGAAGGCCCTTGGGTTGAAAGCTCTTTCAGTCAG 416
Dd	337022	GATCAGCAATGCCCGCTGCTGAAGAAGGCCCTTGGGTTGAAAGACTTTTGTCAAGG 336963
OY	417	AAGAAAAGTTACGGTAAATATATGTAACCATGACCGTATTCAGACAGAAAGAACCCGGC 476
Dd	336962	AAGAAAAGTTGTTGCTAATATTCAGACCGCTGATGACCGTATTCGAAGATTAAGCACCCGC 336903
OY	477	TAACTACGTCGACAGAGCGCGGTAAATACGTAGGGTGAAGGCTTAATTCGAAATTAAC 536
Dd	336902	TAACTACGTCGACAGAGCGCGGTAAATACGTAGGGTGAAGGCTTAATTCGAAATTAAC 336843
OY	537	GCGTAAAGGATGCGCAGCGCGGCTTGAATGTCAGATGTGAATATCCCAGGCTTAAACCTGG 596
Dd	336842	GCGTAAAGGCGCGCAGACGCTTACTTAAGCAGATGTGAATATCCCAGGCGCTCAACCCGG 336783
OY	597	GAATTTGCGTTGAAACTCAAGAGCTAGAGTGTGSCAGAGGAGGTGGAATTCATCTGTGA 656
Dd	336782	GAATTTGCGTTGAAACTCGAGTGTGCTCAGAGGAGGTGGAATTCACAGCTGTGA 336723
OY	657	GCACTGAAATGCGTGAATATATGAAAGAACTCCATGTCGAAAGGACGCTCTCGGGTTAA 716
Dd	336722	GCACTGAAATGCGTGAATATATGAAAGAACTCCATGTCGAAAGGACGCTCTCGGGACAA 336663
OY	717	CACCTGAGCTCATGACAGAAAGCGTGGGAGACAAACAGATTAAGATTAACCTCTGTAGTCCA 776
Dd	336662	CACCTGAGCTCATGACCGAAAGCGTGGGATGCAAAACAGATTAAGATTAACCTCTGTAGTCCA 336603
OY	777	CGCCCTTAACGATGCACTAGTGTGTTGGGCTTTATA--GGCTTGATTAAGAGCTAACG 835

```
Db      336602  CGCCCTAAACGATGTCATTAAGTCGTTGGGCAACCTGATGCTGTAAGCTTAACG 336543
Qy      836    CGTGAAGTGAACGCGCTGGGGAGTACGTCGCAAGATTTAAACTCAAGAGATTAAGACGG 895
Db      336342  CGTAAATTGACCCCTGGGGAGTACGTCGCAAGATTTAAACTCAAGAGATTAAGACGG 336483
Qy      896    GACCCGCAACAAGCGGTGATTAATGATTAATTAATGATCAACCGCAAAAACCTTACCTA 955
Db      336482  GACCCGCAACAAGCGGTGATTAATGATTAATTAATGATCAACCGCAAAAACCTTACCTG 336423
Qy      956    CCTTGAATGATTAAGCAATTTTCTAGATTAATTAAGTCTTGGGCAACCTTAACAG 1014
Db      336422  GTCTTGAATGATTAAGCAATTTTCTAGATTAATTAAGTCTTGGGCAACCTTAAACAG 336363
Qy      1015  GTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074
Db      336362  GTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 336303
Qy      1075  GCAACCCCTTGTCAATTAATGCTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1134
Db      336302  GCAACCCCTTGTCAATTAATGCTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 336243
Qy      1135  AAACCGAGAGAGTGGGGATGAGTCAAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1194
Db      336242  AAGCCGAGAGAGTGGGGATGAGTCAAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 336183
Qy      1195  ACGTAATCAATGCGCGCTGACAGAGGCTGCAACCGCGAGGGGGAGCTATCTACAGA 1254
Db      336182  ACGTAATCAATGCGCGCTGACAGAGGCTGCAACCGCGAGGGGGAGCTATCTACAGA 336123
Qy      1255  AGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1314
Db      336122  AAGCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 336063
Qy      1315  AATGCGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1374
Db      336062  AATGCGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 336003
Qy      1375  CACCATGGAGTGGGGTTTCCACGAGAGGAGTGTCTAAGCTTAAGAGAGGGGCTTGGCA 1434
Db      336002  CACCATGGAGTGGGGTTTCCACGAGAGGAGTGTCTAAGCTTAAGAGAGGGGCTTGGCA 335943
Qy      1435  CGGTGAGATTCATGACTGGGGTG 1457
Db      335942  CGGTGAGATTCATGACTGGGGTG 335920

RESULT 55
AAF21544
ID      AAF21544 standard; DNA; 349980 BP.
XX      AC      AAF21544;
XX      AC      13-MAR-2001 (first entry)
XX      DE      Neisseria meningitidis B nucleotide sequence SEQ ID NO:1.
XX      KM      Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
XX      KM      diagnosis; antigen; detection; infection; gene therapy; antibacterial;
XX      OS      Neisseria meningitidis.
XX      OS      WO200066791-A1.
XX      PD      09-NOV-2000.
XX      PF      08-MAR-2000; 2000WO-US005928.
XX      PR      30-APR-1999; 99US-0132068P.
XX      PR      08-OCT-1999; 99WO-US023573.
XX      PR      28-FEB-2000; 2000GB-00004695.
```

```
XX      PA      (CHIR ) CHIRON CORP.
XX      PA      (GENO-) INST GENOMIC RES.
XX      PI      Piza M, Hickey E, Peterson J, Tettein H, Venter JC,
XX      PI      Meisner V, Galeotti C, Mora M, Ratti G, Scarsella M, Scariato V,
XX      PI      Rappuoli R, Frazer CM, Grandi G;
XX      DR      WPI; 2000-647603/62.
XX      PT      Neisseria meningitidis B full length genome sequence and open reading
XX      PT      frames are used to detect, treat and prevent Neisserial infections.
XX      PS      Claim 7; Appendix A; 692pp; English.
XX      CC      The present invention describes the full length genome of Neisseria
XX      CC      meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
XX      CC      represent fragments of the NMB genomic sequence, as the sequence which
XX      CC      long to go in a record on its own it was split into 8 sequences by 4980 bp
XX      CC      overlap each other at the beginning and end of each sequence by 4980 bp
XX      CC      (i.e. the last 4980 bp of AAF21544 is repeated at the beginning of
XX      CC      AAF21607, the last 4980 bp of AAF21607 are repeated at the beginning of
XX      CC      AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
XX      CC      given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
XX      CC      primers which are used in the exemplification of the present invention.
XX      CC      The NMB genome and fragments from it have antibacterial activity, and can
XX      CC      be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
XX      CC      and/or antibodies which binds to the proteins can be used in compositions
XX      CC      for treating or preventing infection due to Neisserial bacteria or as a
XX      CC      diagnostic reagent for detecting the presence of Neisserial bacteria or
XX      CC      of antibodies raised to Neisserial bacteria. Computers, computer memory,
XX      CC      computer storage medium or computer databases can be used in a search to
XX      CC      identify open reading frames (ORFs) or coding sequences within the NMB
XX      CC      genome. The DNA sequences provide further opportunities to find antigenic
XX      CC      or immunogenic proteins which are more effective in vaccines than the
XX      CC      outer membrane proteins currently used
XX      SQ      Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 U; 0 Other;

Query Match      76.9%; Score 1119.8; DB 3; Length 349980;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 1275; Conservative 0; Mismatches 182; Indels 6; Gaps 4;

Qy      1      ATTGAAGCTGCGCGCATGCTTTACATGGAATCCAGCGGCGAC--GGATGCTTGC 58
Db      307742  ATTGAAGCTGCGCGCATGCTTTACATGGAATCCAGCGGCGACGACGAGAGCTTGC 307801
Qy      59      ATCT--GGTGGCGAGTGGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATG 116
Db      307802  TTCTGGGTGGCGAGTGGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATG 307861
Qy      117      GGTAAAGCATGAAAGATGCTAATACCGCATTAATCTTAAGAGAGGAAACGAGGGATC 176
Db      307862  GATAAAGCATGAAAGATGCTAATACCGCATTAATCTTAAGAGAGGAAACGAGGGATC 307921
Qy      177      GAAAGACCTTGGCTTTGAGCGCGCGATGCTGATTAAGTAACTGTTGGGTGAAAGGC 236
Db      307922  TTGGGGCTTGGCTTTGAGCGCGCGATGCTGATTAAGTAACTGTTGGGTGAAAGGC 307981
Qy      237      CTACCAAGGCGAGCATGATGATGTTGCTGAGAGAGCAGCAGCAGCAGCAGCAGCAGC 296
Db      307982  CTACCAAGGCGAGCATGATGATGTTGCTGAGAGAGCAGCAGCAGCAGCAGCAGCAGC 308041
Qy      297      CACGGCCAGACTCTACGAGAGGAGCAGTGGGGAATTTTGAACAATGGGCGCAAGCTT 356
Db      308042  CACGGCCAGACTCTACGAGAGGAGCAGTGGGGAATTTTGAACAATGGGCGCAAGCTT 308101
Qy      357      GATCCAGCAATGCGCGTGAAGTGAAGAGGCTTTCGGTGTGAAAGCTTTCACTGAG 416
Db      308102  GATCCAGCAATGCGCGTGAAGTGAAGAGGCTTTCGGTGTGAAAGCTTTTGTCAAGG 308161
Qy      417      AAGAAAGGTTACGGTAAATATGCTGATGACCATGACGCTATGCAAGAGAGAGCAGGC 476
```

```
Db 308162 AAGAAAGGCGTGTCTAATATACAGCGGCTGATACGGTACTGGAAGATTAAGCACCGGC 308221
QY 477 TAACTACGTGCGACGACCGCGGTAAATCGTAGGGTGCAGCGTTAATCGGAATTACTCG 536
Db 308222 TAACTACGTGCGACGACCGCGGTAAATCGTAGGGTGCAGCGTTAATCGGAATTACTCG 308281
QY 537 GCGTAAAGGGGCGCAGCGCGCTTGTAAAGTCAAGTGAATATCCCGGGCTTAACCTGG 596
Db 308282 GCGTAAAGGGGCGCAGCGCTTGTAAAGTCAAGTGAATATCCCGGGCTTAACCTGG 308341
QY 597 GAATTCGCTTTGAAATACAAAGCTAGAGTGTGAGAGGAGGAGGTGAATTCATGTGT 656
Db 308342 GAATTCGCTTTGAAATACAAAGCTAGAGTGTGAGAGGAGGAGGTGAATTCATGTGT 308401
QY 657 GCACTGAAATGCGTAGAGATATGGAAGAACATGATGGCGAAGGACGCTCTTGGGTAA 716
Db 308402 GCACTGAAATGCGTAGAGATATGGAAGAACATGATGGCGAAGGACGCTCTTGGGTAA 308461
QY 717 CACTGAGGCTCATGACGAAAGCGTGGGGAGCAAAAGATTAGATACCTGTGATGCA 776
Db 308462 CACTGAGGCTCATGACGAAAGCGTGGGGAGCAAAAGATTAGATACCTGTGATGCA 308521
QY 777 CGCCCTAAAGCATGTCAATAGTGTGGGCTTATTA-GGCTTGTAAAGAACCTAACG 835
Db 308522 CGCCCTAAAGCATGTCAATAGTGTGGGCTTATTA-GGCTTGTAAAGAACCTAACG 308581
QY 836 CGTGAAGTTGACCGCTGGGGAGTACGCTGCGCAAGATTTAAACCTCAAGAAATTGACGG 895
Db 308582 CGTGAAGTTGACCGCTGGGGAGTACGCTGCGCAAGATTTAAACCTCAAGAAATTGACGG 308641
QY 896 GACCCGCAAGAGCGGTGATATATGATTAATTCATGCAACGCAAAACCTTACTTA 955
Db 308642 GACCCGCAAGAGCGGTGATATATGATTAATTCATGCAACGCAAAACCTTACTTA 308701
QY 956 CCCTTGAATGATGACGAATTTTCTAGAGATAGATTAGT-CTTGGGAAAGCGTAAACAG 1014
Db 308702 GTCTTGAATGATGACGAATTTTCTAGAGATAGATTAGT-CTTGGGAAAGCGTAAACAG 308761
QY 1015 GTCTGATGATGCTGTCTGACGTGTCTGTGATGATGTTGGTTAAGTCCCGCAACGAGC 1074
Db 308762 GTCTGATGATGCTGTCTGACGTGTCTGTGATGATGTTGGTTAAGTCCCGCAACGAGC 308821
QY 1075 GCAACCCCTGTCATTAATTTGTCATCATTTGGTGGGACTTAAATGACAGTCCGGTAC 1134
Db 308822 GCAACCCCTGTCATTAATTTGTCATCATTTGGTGGGACTTAAATGACAGTCCGGTAC 308881
QY 1135 AAACCGAGGAGGTGGGATGACGTCAAGTCTCATGACCCTTATGGGTAGGGCTTAC 1194
Db 308882 AAACCGAGGAGGTGGGATGACGTCAAGTCTCATGACCCTTATGGGTAGGGCTTAC 308941
QY 1195 ACGTAAATCAATGCGCGGTACAGAGGGTGGCAACCGCGAGGGGAGCTTAATCTCAAG 1254
Db 308942 ACGTAAATCAATGCGCGGTACAGAGGGTGGCAACCGCGAGGGGAGCTTAATCTCAAG 309001
QY 1255 AGGCGTGTGATGTCGAGTCCGAGTGCACATCGACTCGGTGAAGTCCGAATGCTAGT 1314
Db 309002 AAGCGATGTAGTCCGAGTTCATCTGCAACTGAGTGCATGAAGTCCGAATGCTAGT 309061
QY 1315 AATGCGGATCGACGATGCGCGGTAAATACGTTCCCGGGCTTTGTAAACAACCGCGCTCA 1374
Db 309062 AATGCGGATCGACGATGCGCGGTAAATACGTTCCCGGGCTTTGTAAACAACCGCGCTCA 309121
QY 1375 CACCATGGAGTGGGTTTACACAGAGCAGTACTTAAACCTTAAGAGAGGGCGCTTGGCA 1434
Db 309122 CACCATGGAGTGGGTTTACACAGAGCAGTACTTAAACCTTAAGAGAGGGCGCTTGGCA 309181
QY 1435 CGGTGAGATTCATGACTGGGGTG 1457
Db 309182 CGGTGATTCATGACTGGGGTG 309204
```

RESULT 56  
AAK26285

```
ID AAK26285 standard; DNA; 1477 BP.
XX
AC AAK26285;
XX
DT 25-MAY-1999 (first entry)
XX
DE 16S rDNA sequence of a trichloroethylene-decomposing microbe.
XX
KW Trichloroethylene; decomposition; microbe; Burkholderia; 16S rDNA;
nucleic acid detection; chemical-resistance; ss.
XX
XX Burkholderia sp.
XX
OS Burkholderia sp.
XX
PN JF11056370-A.
XX
PD 02-MAR-1999.
XX
PF 19-AUG-1997; 97JP-00236452.
XX
PR 19-AUG-1997; 97JP-00236452.
XX
PA (TOYT ) TOYOTA JIDOSHA KK.
XX
PA (TOYW ) TOYOTA CHUD KENKUSHO KK.
XX
WP1; 1999-222383/19.
XX
PT Detection of a trichloroethylene-decomposing microbe - using a medium
containing the chemical.
XX
PS Claim 1; Page 11; 17pp; Japanese.
XX
CC The present sequence represents a 16S rDNA of a trichloroethylene-
decomposing microbe, Burkholderia sp. N16-1 (FERM BP-5504). The invention
relates to a method for the detection of a trichloroethylene-decomposing
microbe, Burkholderia sp. N16-1 (FERM BP-5504) in which it is selectively
detected or counted by a medium containing the chemical by using the
CC chemical-resistant strain; it also provides a mutant of the
trichloroethylene-decomposing microbe, Burkholderia sp. N16-1 (FERM BP-
5504) having chemical resistance and maintaining trichloroethylene-
decomposing activity
CC
XX
SQ Sequence 1477 BP; 372 A; 346 C; 474 G; 285 T; 0 U; 0 Other;
XX
Query Match 76.8%; Score 1119.4; DB 2; Length 1477;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 1274; Conservative 0; Mismatches 171; Indels 12; Gaps 4;
QY 6 ACGCTGGCGGATGCTTTACATGCAATGCAAGTGCAGCGCAGACGATGCTTGCATCTGT 65
Db 4 ACGCTGGCGGATGCTTTACATGCAATGCAAGTGCAGCGCAGACGATGCTTGCATCTGT 61
QY 66 GGGAGTGGGAGCGGTGATTAATGATTCGGAACGATTCAGAAAGAGGGGGTTAACGCA 125
Db 62 GGGAGTGGGAGCGGTGATTAATGATTCGGAACGATTCAGAAAGAGGGGGTTAACGCG 121
QY 126 TCGAAGATGTGCTAATACCGCATATACCTTAAGAGAGAAAGAGGGGATC--GAAAGAC 183
Db 122 GCGAAGCCGGATTAATACCGCATATACCGATGTGAGTGAAGAGGGGGCTCTTCGGGAC 181
QY 184 CTTCGCTTTTGGAGCGCGCGATGTCTGATTAGCTTATGTTGGGTAAAGGCTTACCA 243
Db 182 CTCGCGCTAACAGGGGGCGCGCGATGCGAGATTACTGATGGGTAAAGGCTTACCA 241
QY 244 GGGAGCATATAGTATGTTGCTGAGAGAGACCAACCACTCGGACATGAGACAGGCC 303
Db 242 GGGAGCATATGATGCTGTGCTGAGAGAGACCAACCACTCGGACATGAGACAGGCC 301
QY 304 CAGACTCTTACGGGAGGAGCAGTGGGAAATTTTGGCAATGAGCGCAAGCTGATTCAG 363
Db 302 CAAACTCTTACGGGAGGAGCAGTGGGAAATTTTGGCAATGAGCGCAAGCTGATTCAG 361
QY 364 CAATGCGCGTGAAGAGAGCCCTTGGGTTTGAAGCTCTTTGACAGTGAAGAAAA 423
```

```
Db      362 CAATGCCGCGTGTGTGAAGAGCGCTTCGGGTGTGTAAAGCACTTTTGTCCGGAAGAAAA 421
Qy      424 GGTACGGTAAATTAATGTAACCCATGACGGTATCGACAGAAAGAACGCCGGCTAACTAC 483
Db      422 CTCGTGTGTTAAATACCGTGGGGATGACGGTACCGGAAGATTAAGACACGGGCTAACTAC 481
Qy      484 GTGCCAGACGCCGGGTAAATACGTAGGGGTGCAAGCCGTTAATCGGAATTACTGGCGTAAA 543
Db      482 GTGCCAGACGCCGGGTAAATACGTAGGGGTGCAAGCCGTTAATCGGAATTACTGGCGTAAA 541
Qy      544 GGGTGGCCAGCGCCCTTTTAAAGTCAATGTGAATCCCGGGGCTTAAACCTGGGAATTCG 603
Db      542 GCGGCGCCAGCGCGGTTCGCTAAAGACAGATGTGAATCCCGGGCTTAAACCTGGGAATTCG 601
Qy      604 GTTTGAAGACTACAAAGCTAGAGTGTGGCAGAGAGGTGGAATTCATGTGTACAGTGA 663
Db      602 ATTTGTAACTGTGGCGGCTTAAGATATGGCAGAGGGGGTAAAGATTCCAGTGTACAGTGA 661
Qy      664 AATGCCGTAGAGATATGGAAGAACATCGATGCGAAGCGACCTTCCTGGGTTAACACTGAC 723
Db      662 AATGCCGTAAAGATGTGAGAGAAATACCGATGGCGAAGCGACGCCCTGGGCGAATATCTGAC 721
Qy      724 GCTATATGACAGAAAGCGTGGGAGACAAACAGGATTAATACCTCGTATGTCACAGCCCTA 783
Db      722 GCTCATGACAGAAAGCGTGGGAGACAAACAGGATTAATGATACCTCGTATGTCACAGCCCTA 781
Qy      784 AAGCATGCAACTAGTTGTGGGCTTATAGGCTGTGTAACGAAGCTAACGCGCTGAAGT 843
Db      782 AACCATGCAACTAGTTGTGGGATTCATTTCTTATGTAACGTAAAGCTTAACGCGCTGAAGT 841
Qy      844 TGACCGGCTGGGGAGTACGGTTCGCAAGATTAAACTCAAGAGAAATGACGGGAGACCCGCA 903
Db      842 TGACCGGCTGGGGAGTACGGTTCGCAAGATTAAACTCAAGAGAAATGACGGGAGACCCGCA 901
Qy      904 CAAGCGGTGATTAATGTGGAATTAATTCATGCAACGCGAAACCTTAACCTACCTTGAC 963
Db      902 CAAGCGGTGATTAATGTGGAATTAATTCATGCAACGCGAAACCTTAACCTACCTTGAC 961
Qy      964 ATGTAGGGAATTTCTAGAGATAGATTAGTGC---TTGGGGAACGCTAACACAGAGTGTG 1020
Db      962 ATGTATGGAACCTGAGCTGAGAGGTCCCGAAGAGGACATTAACACAGTGTG 1021
Qy      1021 CATGCGTGTGCTGACGCTGCTGTGTGTGAGATGTGGGTTAAGTCCCGCAACGAGCGCAAC 1080
Db      1022 CATGCGTGTGCTGACGCTGCTGTGTGTGAGATGTGGGTTAAGTCCCGCAACGAGCGCAAC 1081
Qy      1081 CTTGTCAATTAATTTGTCATCAATTTGTTGGGCACTTTAATGAGATGCGCGGTGAACAAAC 1140
Db      1082 CTTGTCCCTAGTTGCTAC-----GCMAAGACACTCTAGGAGACCTGCGGTGAACAAAC 1136
Qy      1141 GAGGAAGGTGGGGATGAGTCAAGTCTCATGGCCCTTAATGGGTAGGGCTTCACACGTAA 1200
Db      1137 GAGGAAGGTGGGGATGAGTCAAGTCTCATGGCCCTTAATGGGTAGGGCTTCACACGTAA 1196
Qy      1201 TACAATGCGCGGTACAGAGGGTTGGCAACCCGCGAGGGGAGACTAATCTCAGAAAGCGCG 1260
Db      1197 TACAATGCGCGGTACAGAGGGTTCGCAACCCGCGAGGGGAGCCAAATCCAGAAAAACCGA 1256
Qy      1261 TCGTAGTCCGGAATCGGAAGTCTGCAACTGCACTCCGTGAAGTCCGAATGCTAAGTAAATGC 1320
Db      1257 TCGTAGTCCGGAATCGCACTGCAACTGCAAGTGCCTGAAGCTGGAATGTTAGTAAATGC 1316
Qy      1321 GGATCAGATGCGCGGTGAATAGTTCGCGGGCTTTTGAACACACGCCCGTCAACACCAT 1380
Db      1317 GGATCAGATGCGCGGTGAATAGTTCGCGGGCTTTTGAACACACGCCCGTCAACACCAT 1376
Qy      1381 GGGAGTGGGTTTCAACAGAAAGAGGTAGTCTAACCGTAAAGAGGGCGCTTGCAAGGTGA 1440
Db      1377 GGGAGTGGGTTTCAACAGAAAGTGGTAAAGTCTAACCGCAAGAGGACGCGTCAACACGTTAG 1436
Qy      1441 GATTATGACTGGGGT 1457
Db      1437 GATTATGACTGGGGT 1453
```

```
RESULT 57
ADB61691
ID ADB61691 standard; DNA; 1610 BP.
XX
AC ADB61691;
XX
DT 04-DEC-2003 (first entry)
XX
DE 16S rRNA of Burkholderia pseudomallei DNA sequence.
XX
KW enriching mRNA; high quality bacterial mRNA; bacterial gene expression;
KW poly-A tail; mRNA purification; oligo-dT capture;
KW prokaryote mRNA purification; bridging oligonucleotide; targeting region;
KW capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;
KW eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;
KW 28S eukaryotic RNA bridging oligonucleotide; ds.
XX
OS Burkholderia pseudomallei.
XX
PN W02003054162-A2.
XX
PD 03-JUL-2003.
XX
PF 19-DEC-2002; 2002MO-US041014.
XX
PR 20-DEC-2001; 2001US-00029397.
XX
PA (AMBI-) AMBION INC.
XX
PI Murphy GL, Whitley JP;
XX
DR WPI; 2003-663255/62.
XX
PT Depleting or isolating targeted nucleic acids e.g. rRNA, involves using a
PT bridging oligonucleotide comprising bridging region and a targeting
PT region complementary to a targeted nucleic acid, and a capture
PT oligonucleotide.
XX
PS Claim 4; Page 175; 208pp; English.
XX
CC This invention relates to a novel method for isolating, depleting or
CC separating a targeted nucleic acid, such as rRNA, from a sample
CC comprising targeted and non-targeted nucleic acids. It effects a way of
CC enriching for non-targeted nucleic acids such as mRNAs. Isolating
CC sufficient quantities of high quality bacterial mRNA is a demanding
CC process which impedes analysis of bacterial gene expression in the
CC presence of host cells. A small percentage of bacterial mRNAs may be poly
CC -A tailed, but these are targeted for degradation and tend to be
CC unstable. As a result, the commonly employed method for mRNA purification
CC with eukaryotic cells, oligo-dT capture, is ineffective. The present
CC invention provides an alternative, more suitable method for mRNA
CC purification from prokaryotes. The method of the invention comprises the
CC purification of a sample with a bridging oligonucleotide (containing a
CC targeting region) and subsequently incubating with a capture
CC oligonucleotide allowing the isolation of the target from the sample. The
CC method is useful for depleting or isolating targeted nucleic acid, for
CC example rRNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S
CC or 18S or eukaryotic 28S rRNA, from a sample. The rRNA sequence may
CC comprise any one of 64 fully defined sequences as given in the
CC specification. The present sequence is that of a DNA sequence which
CC represents the sequence of 16S rRNA of Burkholderia pseudomallei related
CC to the invention.
XX
SQ Sequence 1610 BP; 394 A; 383 C; 514 G; 319 T; 0 U; 0 Other;
XX
Query Match 76.8%; Score 1118.8; DB 10; Length 1610;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 182; Indels 10; Gaps 3;
Qy 1 ATTGAAGCGTGGCGGATGCTTTACATGCAAGTCAAGCGGACGACGATGCTTCAT 60
|||||
```



Db 82 ATTGAACGCTGGGGGATGCTTACACATGCAAGTGAACGGAGACGCG--GCTTCGGC 139  
 Qy 61 CTGGTGGCGAGTGGCGGACGGGTGAGTATGATCGGAACGTAATCCAGAAAGAGGGGGTA 120  
 Db 140 CTGGTGGCGAGTGGCGGACGGGTGAGTATGATCGGAACGATGCTGTAGTGGGGGATA 199  
 Qy 121 ACCGATCGAAAGATGTGCTAATACCGGATATACCTTAAGAGGAAAAGCAGGGGATGAAA 180  
 Db 200 GCCCGGAGAAAGCGAATTAAATCCGATACGATCTGAGGATGAAGCGGGGACCTTCG 259  
 Qy 181 GACCTTGCGCTTTTGGAGCGCGGATGTCTGATTAGCTAGTGTGGTGAAGGCTTAC 240  
 Db 260 GGGCTGCGCTTAAGGTTGGCCGATGCTGATTAGTATGTGGGGTAAAGGCTTAC 319  
 Qy 241 CAAGGCGACGATGATGTTGCTGAGAGGACGACGACCACTGGGACTGAGACAG 300  
 Db 320 CAAGGCGACGATGATGTTGCTGAGAGGACGACGACCACTGGGACTGAGACAG 379  
 Qy 301 GCCCAAGCTCTTACGGGAGGACAGCTGGGGAAATTTTGGCAATGGGCGGACCTGATC 360  
 Db 380 GCCCAAGCTCTTACGGGAGGACAGCTGGGGAAATTTTGGCAATGGGCGGACCTGATC 439  
 Qy 361 CAGCAATGCCGCTGATGAGAAAGGCTTGGGTTGTAAGCTCTTTCAGTGGAGAGA 420  
 Db 440 CAGCAATGCCGCTGATGAGAAAGGCTTGGGTTGTAAGCTCTTTCAGTGGAGAGA 499  
 Qy 421 AAAGGTTACGGTAATATATCTGACCCATGACGGTATCGACAGAAAGACACGGCTTAC 480  
 Db 500 AATCATTTCTGGCTAATATCCGAGTGTATGACGGTACCGGAATAATTAACACGGGCTTAC 559  
 Qy 481 TAGTGCCAGACACCGCGGTAAATACGTAAGGTGCAAGCGTTAATCCGAATTAATCTGGGCGT 540  
 Db 560 TAGTGCCAGACACCGCGGTAAATACGTAAGGTGCAAGCGTTAATCCGAATTAATCTGGGCGT 619  
 Qy 541 AAAGGTTACGGTAATATATCTGACCCATGACGGTATCGACAGAAAGACACGGCTTAC 600  
 Db 620 AAAGGTTACGGTAATATATCTGACCCATGACGGTATCGACAGAAAGACACGGGCTTAC 679  
 Qy 601 TGGGTTTGAACCTAAGAGTGTGACAGAGGAGGTGAAATTCATGTGTAGCAG 660  
 Db 680 TGCATTGTGACTGTGACAGGTAGTGTGACAGAGGAGGTGAAATTCATGTGTAGCAG 739  
 Qy 661 TGAATGTGCTGATGATGTGAAAGCACTGATGCGGAGGACGCTCTGGGTTAATCACT 720  
 Db 740 TGAATGTGCTGATGATGTGAAAGCACTGATGCGGAGGACGCTCTGGGTTAATCACT 799  
 Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGAGATTAGTACCTGTGTAGTCCAGCC 780  
 Db 800 GACGCTCATGACGAAAGCGTGGGAGCAAAAGAGATTAGTACCTGTGTAGTCCAGCC 859  
 Qy 781 CTAACGATGTCAAATGTTGTGGGCTTATTAAGCTTGTGTAACGAACTTAACGCGTGA 840  
 Db 860 CTAACGATGTCAAATGTTGTGGGCTTATTAAGCTTGTGTAACGAACTTAACGCGTGA 919  
 Qy 841 AGTTGACCGCTGGGAGTACGCTCGGAGATTAATTAAGAGATTGAACGGGGAGCC 900  
 Db 920 AGTTGACCGCTGGGAGTACGCTCGGAGATTAATTAAGAGATTGAACGGGGAGCC 979  
 Qy 901 GCAAGCGGATGATTAATGATTAATTAATGATGATTAATTAATTAATTAATTAATTAATTAAT 960  
 Db 980 GCAAGCGGATGATTAATGATTAATTAATGATGATTAATTAATTAATTAATTAATTAATTAAT 1039  
 Qy 961 GACATGATGAAATTTTCTAGAGATAGATTAGTCT--TCGGGACGCTTAACACAGGTG 1017  
 Db 1040 GACATGATGAAATTTTCTAGAGATAGATTAGTCT--TCGGGACGCTTAACACAGGTG 1099  
 Qy 1018 CTGATGATGCTGTGCTGATGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1077  
 Db 1100 CTGATGATGCTGTGCTGATGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1159  
 Qy 1078 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
 Db 1160 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1214

Qy 1138 CCGGAGGAGAGTGGGGATGACGCTCAAGTCTTCATGAGCCCTTAATGGGTAGGGCTTCACAG 1197  
 Db 1215 CCGGAGGAGAGTGGGGATGACGCTCAAGTCTTCATGAGCCCTTAATGGGTAGGGCTTCACAG 1274  
 Qy 1198 TAATCAATGAGCGCTGACAGAGGAGTGGCCAAACCGCGAGGGGAGCTTAATTCAGAAAC 1257  
 Db 1275 TCATCAATGAGCGCTGACAGAGGAGTGGCCAAACCGCGAGGGGAGCTTAATTCAGAAAC 1334  
 Qy 1258 GCGTGTAGTCCGATCGAGTCTGCAATCGAATCCGTAAGTGGGAATGCTGTAT 1317  
 Db 1335 CGATCTAGTCCGATCGAGTCTGCAATCGAATCCGTAAGTGGGAATGCTGTAT 1394  
 Qy 1318 CCGGATGACAGTGTGCGGCTGATTAATGCTTCCGGCTTGTACACACCGCGTCAAC 1377  
 Db 1395 CCGGATGACAGTGTGCGGCTGATTAATGCTTCCGGCTTGTACACACCGCGCTCAAC 1454  
 Qy 1378 CATGGAGTGGGTTTACAGAGAGAGTGTCTTAACCTTAAGAGAGGCGCTTGGCCAGG 1437  
 Db 1455 CATGGAGTGGGTTTACAGAGAGAGTGTCTTAACCTTAAGAGAGGCGCTTGGCCAGG 1514  
 Qy 1438 TGAGATTCAATGACTGGGTTG 1457  
 Db 1515 TAGGATTCAATGACTGGGTTG 1534

RESULT 58

AB269297  
 ID AB269297 standard; DNA; 1400 BP.

XX AB269297;

DT 11-AUG-2003 (first entry)

XX J lividum 16s ribosomal RNA gene fragment #1.

KM Osteoarthritis; antibacterial; Janthinobacterium; 16s RNA; gene;  
 KW osteopathic; antiarthritic; analgesic; anti-inflammatory; ds.

XX Janthinobacterium lividum.

PN W02002102384-A1.

XX 27-DEC-2002.

PF 17-JUN-2002; 2002WO-GB002771.

PF 15-JUN-2001; 2001GB-00014672.

PR (ORTH-) ORTHOGENICS AS.

PA (GARD/) GARDNER R.

XX E1-Gewely MR;

XX WPI; 2003-17519/17.

PT Use of an antibacterial agent in the manufacture of a medicament for  
 treating osteoarthritis.

PS Example 1; Page 46-48; 899p; English.

CC The present invention relates to the use of an antibacterial agent in the  
 CC production of a treatment for osteoarthritis. The bacteria causing  
 CC osteoarthritis is Janthinobacterium lividum. The present sequence is a  
 CC fragment of the 16s RNA coding sequence from J. lividum shown in the  
 CC exemplification of the invention

SO Sequence 1400 BP; 370 A; 309 C; 429 G; 287 T; 0 U; 5 Other;

Query Match 76.8%; Score 1118.4; DB 8; Length 1400;  
 Best Local Similarity 88.6%; Pred. No. 0;  
 Matches 1246; Conservative 0; Mismatches 151; Indels 10; Gaps 3;



Oy	27	CATGCAGTGCAAACGGCAGCAGCGAATCGTTCGATCTGCTGGCGAGGTGGCGAGCGGGTGAAG	86
Db	1	CATCAAGTCCAAACGGCAGCAGCGA - GCTTGCCTCTGTGTGCGAGGTGGCGAAGCGGGTGAAG	58
Oy	87	TAATGCATCCGAAAGGTATCCAGAGAGAGGGGGGTAAACCATCGAAGAGATGTGCTAATAACCG	146
Oy	147	CATATACTCTAAGAGAGAAAGCAGGGGATCGAAAGACCTTTCGCTTTTGGACGGCCGAT	206
Db	119	CATACGATCTAAGAGATGAAGATGGGGGATCGCAAGACTCATGCTCGTGGACGGCCGAT	178
Oy	207	GTCGATTAAGCTAAGTTGGTGGGGTAAAGGCGCTACCAAGCGGACGATATAGATAGTTGGCTG	266
Db	179	ATCTGATTAAGCTAAGTTGGTGGGTAAAGCTTACCAAGGACATCGATAGTAGTGGTCTG	238
Oy	267	AGAGAGCAGCAGGACCACTGGGACTGAGACACGGCCACAGCTCTTACGGGAGCAGCAG	326
Db	239	AGAGAGCAGCAGGACCACTGGAACCTGAGACACGGTCCAGACTCTTACGGGAGCAGCAG	298
Oy	327	TGGGGAATTTTGGACAATGGGCGGACAGCCTGATCAGCAATGCCGCTGAGTGAAGAAG	386
Db	299	TGGGGAATTTTGGACAATGGGCGGAAGCTGATCAGGAATGCCGCTGAGTGAAGAAG	358
Oy	387	CCTTCGGGTTGAAAGCTCTTTCAGTGCAGAGAAAGGTTACGGTAAATATATCTGAC	446
Db	359	CCTTCGGGTTGAAAGCTCTTTCAGAGGAAAGGACGGTGAAGACTAATATCTTTCGCT	418
Oy	447	CATGACGGTATCGACAGAAAGACACGGGCTATCATAGTGCACACAGCCGCGTAAATAC	506
Db	419	AATGACGGTACTGAAGATTAAGCACCGGCTATCTAGTGCACACAGCCGCGTAAATAC	478
Oy	507	TAGGTTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGGTTGCGACAGCGGCTTGTAG	566
Db	479	TAGGTTGCGAACGTTAATCGGAATTACTGGGCGTAAAGCGTGAAGCGGCTTGTAG	538
Oy	567	TCAGATGTGAATCCCGGGGCTTAACTCGGGAATTTGGCTTTGAAACTTCAAGAGCTAAGT	626
Db	539	TCTGATGTGAATCCCGGGGCTCAACCTGGGAATTTGGAACCTGCAAGGCTTAAGT	598
Oy	627	GTGGCAGAGGGAGGTGGAATTCATGTGTAGCAGTGAATTCGTAGAGATATGGAAGAAC	686
Db	599	CTGGCAGAGGGGGGTAGAAATTCACGTGTAGCAGTGAATTCGTAGATATGGAAGAAC	658
Oy	687	ATCGATGCGCAAGGACGCGCTCTGGGTTAACACTGACCGCTATGCACGAAACGCTGGGGA	746
Db	659	ACCGATGCGAAGGACGCGCCCTGGGTTCAAGTTGACGCTATGACGAAACGCTGGGGA	718
Oy	747	GCAAAACAGATTAAGATTAACCTGTGTATGTCACGCGCTTAAACGATGTCAACTATGTTGTGGG	806
Db	719	GCAAAACAGATTAAGATTAACCTGTGTATGTCACGCGCTTAAACGATGTCTATGTTGTGGG	778
Oy	807	CCTTATTAAGCTTTGGTAAACGAGCTTAAGCGCGGAATGACCGCGCTGGGGAGTACGGTGG	866
Db	779	TCTTATTAAGCTTTGGTAAACGAGCTTAAGCGCGGAATGACCGCGCTGGGGAGTACGGTGG	838
Oy	867	CAAGATTTAAACTCAAGGAATTTGACGGGGACCCGCAAGCGGTGATTAATGTGATTA	926
Db	839	CAGATTTAAACTCAAGGAATTTGACGGGGACCCGCAAGCGGTGATTAATGTGATTA	898
Oy	927	ATTGCATGCAAGCGGAAAACTTACCTTACCTTGCATGTAGCAATTTTCTTAGAGATA	986
Db	899	ATTGCATGCAAGCGGAAAACTTACCTTACCTTGCATGTAGCAATTTCTTAGAGATC	958
Oy	987	GATTAAGTCT---TCGGGAAGCTTAACAAGGTGTGCATGGCTGTGTCAGCTCGTCTC	1043
Db	959	AAGGAGTGTCTCAAAAGAGAACATGTAACAAGGTGTGCATGGCTGTGTCAGCTCGTCTC	1018
Oy	1044	GTGAGATGTTGGGTTAAGTCCCGCAACAGCGCAACCTTGTGCATTAATTCACATATT	1103
Db	1019	GTGAGATGTTGGGTTAAGTCCCGCAACAGCGCAACCTTGTGCATTAATTCAGTTCG-----TAC	1073
Oy	1104	GGTGGGCACTTTAATAGACTGCGGGTGAACAACCGAGGAAGGTGGGATGACGTGAA	1163

[illegible]

CC supplement in livestock and poultry food composition. Note: The present  
CC sequence is not shown in the specification but has been accessed from  
CC Genbank using the appropriate accession number given in the  
CC specification. (Updated on 06-AUG-2003 to correct OS field.)  
XX

SO Sequence 1544 BP; 380 A; 357 C; 492 G; 315 T; 0 U; 0 Other;

Query Match 76.7%; Score 1118.2; DB 5; Length 1544;  
Best Local Similarity 87.1%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 183; Indels 6; Gaps 4;

```
QY 1 ATTGAAGCTGGCGGCGATCTTACACATGCAATGCAAGCGGAGCAC--GATGCTTGC 58
DB 29 ATTGAAGCTGGCGGCGATCTTACACATGCAATGCAAGCGGAGCACAGGAGAGCTTGC 88
QY 59 ATCT--GGTGGCGAGTGGCGGAGCGGCTGAATTCATCGAAAGTATCCAGAAAGAGG 116
DB 89 TTCCTGGGTGGCGAGTGGCGGAGCGGCTGAATTCATTCGAAAGTATCCGGGTAGCGGG 148
QY 117 GGTAAAGCATCGAAAGTGTCTTAATCCGCAATTAATCTTAAGAGAGAAAGCAGGGATC 176
DB 149 GATTAAGCATCGAAAGTGTCTTAATCCGCAATTAATCTTAAGAGAGAAAGCAGGGATC 208
QY 177 GAAGAAGCTTGGCGGCTTGGAGCGGCGGATGCTGATTAAGCTAGTGGGTAAAGGC 236
DB 209 TTGGGGCTTGGCGGCTTGGAGCGGCGGATGCTGATTAAGCTAGTGGGTAAAGGC 268
QY 237 CTACCAAGGCGAGCATGATGTTGCTTGAAGAGAGCAGCAGCAGCATGCGAGCTGAGA 296
DB 269 CCACCAAGGCGAGCATGATGTTGCTTGAAGAGAGCAGCAGCAGCATGCGAGCTGAGA 328
QY 297 CAGGCGCCAGATCTTACCGGAGAGCAGCATGCGGAGATTTTGAACAATGGGCCAAGCCT 356
DB 329 CAGGCGCCAGATCTTACCGGAGAGCAGCATGCGGAGATTTTGAACAATGGGCCAAGCCT 388
QY 357 GATCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTGTAAGCTCTTTCAGTGCAG 416
DB 389 GATCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTGTAAGCTCTTTCAGTGCAG 448
QY 417 AAGAAAGGTTACGGTAATTAATCTGTGACCCATGACCGGATTCAGACAGAAAGACCGGC 476
DB 449 AAGAAAGGTTACGGTAATTAATCTGTGACCCATGACCGGATTCAGACAGAAAGACCGGC 508
QY 477 TAACTAAGTCCAGCAGCGCGGCTTAATCTAAGGTGCAAGCGTAACTCGAATTAATCTG 536
DB 509 TAACTAAGTCCAGCAGCGCGGCTTAATCTAAGGTGCAAGCGTAACTCGAATTAATCTG 568
QY 537 GCGTAAAGGTTGCGCAGCGCGGCTTGAAGTCAAGTGAATCCCGGCTTAACTCG 596
DB 569 GCGTAAAGGTTGCGCAGCGCGGCTTGAAGTCAAGTGAATCCCGGCTTAACTCG 628
QY 597 GAAATGGCTTGAATTAATCAAGAGTGAAGTGGCAGAGGAGGTGAATTCATGTTGA 656
DB 629 GAAATGGCTTGAATTAATCAAGAGTGAAGTGGCAGAGGAGGTGAATTCATGTTGA 688
QY 657 GCAGTGAATTCGTAAGATTAAGAAACAATCGATGCGAAGCGCGCTTCGAGGTAA 716
DB 689 GCAGTGAATTCGTAAGATTAAGAAACAATCGATGCGAAGCGCGCTTCGAGGTAA 748
QY 717 CACTGAAGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGTACCTCGTATGCA 776
DB 749 CACTGAAGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGTACCTCGTATGCA 808
QY 777 CGCCCTAAAGCATGCTAATAGTTGTTGGCTTATTA--GGCTTGTAAAGAAAGTAAAG 835
DB 809 CGCCCTAAAGCATGCTAATAGTTGTTGGCTTATTA--GGCTTGTAAAGAAAGTAAAG 868
QY 836 CGTGAAGTTAAGCGGCTGGGAGAGTACGCTGCAAGATTAAATCTCAAAAGAAATTGACGG 895
DB 869 CGTGAAGTTAAGCGGCTGGGAGAGTACGCTGCAAGATTAAATCTCAAAAGAAATTGACGG 928
QY 896 GATCCGCAAGAGGAGTGAATTAATGATGATTAATTCAGTGAAGCGGAAACCTTACCTA 955
```

```
DB 929 GACCCGACAAAGCGGTGATGATGATTAATTCAGTCAACGCGAAGAACTTACTCG 988
QY 956 CCCTTGACATGTAAGCAATTTTCTAGATTAATTAATG--CTTCGGGAAGCTTAACAG 1014
DB 989 GTTTGACATGTCGGAATCTCCGAGACGAGAGAGTCTTCGGAGCCGTAACAG 1048
QY 1015 GTGCTGATGCTGTCGTCAGCTGTCGTGAGATGTTGGTTAAAGTCCCGCAAGAC 1074
DB 1049 GTGCTGATGCTGTCGTCAGCTGTCGTGAGATGTTGGTTAAAGTCCCGCAAGAC 1108
QY 1075 GCAACCTTGTCAATTAATTCATCAATCTTGGTGGCACTTAAATGAGACTGCGGTAC 1134
DB 1109 GCAACCTTGTCAATTAATTCATCAATCTTGGTGGCACTTAAATGAGACTGCGGTAC 1168
QY 1135 AAACCGAAGAGTGGGAGATACGTCAGTCTCAATGGCCCTTAATGGGTAGGGCTTAC 1194
DB 1169 AAACCGAAGAGTGGGAGATACGTCAGTCTCAATGGCCCTTAATGAGAGGCTTAC 1228
QY 1195 ACGTAATCAATGAGCGCTACAGAGGTTGCCAACCCCGGAGGGAGCTAAATTCAGAA 1254
DB 1229 ACGTCAATCAATGTCGTGTCAGAGGTTGCCAACCCCGGAGGGAGCTAAATTCAGAA 1288
QY 1255 ACGCGCTGTAGTCCGATCGAGTGTGCACTGCACTCCGTAAAGTCCGAATTCGTAGT 1314
DB 1289 AACCGATCGTATGTCGAGTTGCACTTCCATCTGCACTGCACTGCAATTCGTAGT 1348
QY 1315 AATCGCGATCGATCGAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1374
DB 1349 AATCGCGATCGATCGAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1408
QY 1375 CACCATGAGAGTGGGTTCACCAAGAGCAGTGTCTTAACCGTAAAGAGAGGCGCTTGCA 1434
DB 1409 CACCATGAGAGTGGGTTCACCAAGAGTGTCTTAACCGTAAAGAGAGGCGCTTGCA 1468
QY 1435 CCGTGAATTCATGACTGGGGTG 1457
DB 1469 CCGTGAATTCATGACTGGGGTG 1491

RESULT 60
ADB61692
ID ADB61692 standard; DNA; 1544 BP.
XX
XX
AC ADB61692:
XX
DT 04-DEC-2003 (first entry)
XX
DE 16S rRNA of Neisseria gonorrhoeae DNA sequence.
XX
XX enriching mRNA; high quality bacterial mRNA; bacterial gene expression;
XX poly-A tail; mRNA purification; oligo-dT capture;
XX prokaryotic mRNA purification; bridging oligonucleotide; targeting region;
XX capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;
XX eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;
XX 28S eukaryotic rRNA bridging oligonucleotide; db.
XX
OS Neisseria gonorrhoeae.
XX
PN MO2003054162-A2.
XX
PD 03-JUL-2003.
XX
PF 19-DEC-2002; 2002MO-US041014.
XX
PR 20-DEC-2001; 2001US-00029397.
XX
PA (AMBI-) AMBION INC.
XX
PI Murphy GL, Whitley JP;
XX
XX MPI; 2003-663255/62.
XX
PT Depleting or isolating targeted nucleic acids e.g. rRNA, involves using a
```

PT bridging oligonucleotide comprising bridging region and a targeting  
PT region complementary to a targeted nucleic acid, and a capture  
PT oligonucleotide.

PS Claim 4, Page 175-176; 208bp; English.

XX This invention relates to a novel method for isolating, depleting or  
CC separating a targeted nucleic acid, such as rRNA, from a sample  
CC comprising targeted and non-targeted nucleic acids. It effects a way of  
CC enriching for non-targeted nucleic acids such as mRNAs. Isolating  
CC sufficient quantities of high quality bacterial mRNA is a demanding  
CC process which impedes analysis of bacterial gene expression in the  
CC presence of host cells. A small percentage of bacterial mRNAs may be poly  
CC -A tailed, but these are targeted for degradation and tend to be  
CC unstable. As a result, the commonly employed method for mRNA purification  
CC with eukaryotic cells, oligo-dT capture, is ineffective. The present  
CC invention provides an alternative, more suitable method for mRNA  
CC purification from prokaryotes. The method of the invention comprises the  
CC incubation of a sample with a bridging oligonucleotide (containing a  
CC targeting region) and subsequently incubating with a capture  
CC oligonucleotide allowing the isolation of the target from the sample. The  
CC method is useful for depleting or isolating targeted nucleic acid, for  
CC example rRNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S  
CC or 18S or eukaryotic 28S rRNA, from a sample. The rRNA sequence may  
CC comprise any one of 64 fully defined sequences as given in the  
CC specification. The present sequence is that of a DNA sequence which  
CC represents the sequence of 16S rRNA of *Neisseria gonorrhoeae* related to  
CC the invention.

XX Sequence 1544 BP; 380 A; 357 C; 492 G; 315 T; 0 U; 0 Other;

Query Match Best Local Similarity 76.7%; Score 1118.2; DB 10; Length 1544;

Matches 1274; Conservative 0; Mismatches 183; Indels 6; Gaps 4;

QY	1	ATTGAACGCTGCGCGCATGCTTTTACATGCAATGCGAAGCGGAGAC--GATGCTTCG	58
DB	29	ATTGAACGCTGCGCGCATGCTTTTACATGCAATGCGAAGCGGAGAC--GATGCTTCG	88
QY	59	ATCT--GGTGGCGAGTGGCGGAGCGGTGATGCAATGCGAAGCGTATCCAGAAAGGAG	116
DB	89	TTCTCGGAGTGGCGGAGTGGCGGAGGTGATGCAATGCGAAGCGTATCCAGAAAGGAG	148
QY	117	GATACGCGATCGAAGATGCTGCTAATACCGCATATATCTAAGAGAGAAAGCGGAGATC	176
DB	149	GATACGCGATCGAAGATGCTGCTAATACCGCATATATCTGTAAGAGAGAAAGCGGAGATC	208
QY	177	GAAAGACCTTGGCTTTTGAAGCGGCGGATGCTGATTAAGCTAGTTGGTGGGTTAAAGC	236
DB	209	TTGGGCGCTTGGCTTATCGAAGCGGCGGATGCTGATTAAGCTAGTTGGTGGGTTAAAGC	268
QY	237	CTACCAAGCGGCGGATGCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	296
DB	269	CCACCAAGCGGCGGATGCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	328
QY	297	CACGCGCGGAGACTCTTACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT	356
DB	329	CACGCGCGGAGACTCTTACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT	388
QY	357	GATTCAGCAATGCGCGGTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	416
DB	389	GATTCAGCAATGCGCGGTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	448
QY	417	AAGAAAGGTTACGTAATATATCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	476
DB	449	AAGAAAGGTTACGTAATATATCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	508
QY	477	TAACTACGTCGAG	536
DB	509	TAACTACGTCGAG	568
QY	537	GGCTAAAGGAGTGGAG	596

DB	569	GGCTAAAGGAGTGGAG	628
QY	597	GAAATGGCTTTGAAACTTACAAAGCTAGAGTGGCGAGAGAGAGAGAGAGAGAGAGAG	656
DB	629	GAACTGGCTTTGAAACTTACAAAGCTAGAGTGGCGAGAGAGAGAGAGAGAGAGAGAG	688
QY	657	GCAGTGAATGCGGTAGAGATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	716
DB	689	GCAGTGAATGCGGTAGAGATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	748
QY	717	CACGAGCGCTCATGCAAGAAAGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	776
DB	749	CACGAGCGCTCATGCTCGAAGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	808
QY	777	CGCCCTTAACGATGTCATCTAGTTGTTGGGCTTATTA--GGCTTGGTAAGAGAGAGAG	835
DB	809	CGCCCTTAACGATGTCATCTAGTTGTTGGGCTTATTA--GGCTTGGTAAGAGAGAGAG	868
QY	836	CGTGAAGTTGACCGCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	895
DB	869	CGTGAAGTTGACCGCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	928
QY	896	GACCCGCAAGAGCGGTGATTAATGATTAATTCATGCAAGCGGAAAGAGAGAGAGAGAG	955
DB	929	GACCCGCAAGAGCGGTGATTAATGATTAATTCATGCAAGCGGAAAGAGAGAGAGAGAG	988
QY	956	CCCTTGACATGTAAGCAATTTTCTAGAGATGATTAAGT--CTTCGGGAGAGAGAGAGAG	1014
DB	989	GTTTTGACATGTAAGCAATTTTCTAGAGATGATTAAGT--CTTCGGGAGAGAGAGAGAG	1048
QY	1015	GTGCTGATGAGCGCTGCTGCAAGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG	1074
DB	1049	GTGCTGATGAGCGCTGCTGCAAGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG	1108
QY	1075	GCAACCGCTTGCATTAATGCGCATCATTTGTTGGGCACTTTAATGAGAGAGAGAGAGAG	1134
DB	1109	GCAACCGCTTGCATTAATGCGCATCATTTGTTGGGCACTTTAATGAGAGAGAGAGAGAG	1168
QY	1135	AAACCGAG	1194
DB	1169	AAACCGAG	1228
QY	1195	ACGTAATACATGAGCGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1254
DB	1229	ACGTAATACATGAGCGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1288
QY	1255	AGCGGTGATGATGCGGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1314
DB	1289	AACGATGATGATGCGGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1348
QY	1315	AATGCGGATGAGATGCTGCGGTGAATACGTTCCCGGCTTTGTAACACAGAGAGAGAGAG	1374
DB	1349	AATGCGGATGAGATGCTGCGGTGAATACGTTCCCGGCTTTGTAACACAGAGAGAGAGAG	1408
QY	1375	CACCATGAG	1434
DB	1409	CACCATGAG	1468
QY	1435	CGGTGAGATGATGAG	1491
DB	1469	CGGTGAGATGATGAG	1525

RESULT 61  
AA18765  
ID AA18765 standard; rRNA; 1536 BP.  
XX  
XX AA18765;  
AC  
XX  
XX 16-OCT-2003 (revised)  
DT 05-JUL-1996 (first entry)  
XX  
XX Pseudomonas testobacterium 16S ribosomal RNA.

XX Atrazine; pesticide degradation; soil decontamination; bioremediation;  
KW s-triazine; herbicide; ss.  
XX Comamonas testosteroni.  
OS  
XX  
FH Key Location/Qualifiers  
FT misc\_difference 46  
FT /\*tag= a  
FT /note= "base 46 is identified as 'n'"  
PN USS508193-A.  
XX  
XX 16-APR-1996.  
PD 31-AUG-1993; 93US-00114695.  
XX 31-AUG-1993; 93US-00114695.  
PR 31-AUG-1993; 93US-00114695.  
XX (MINU ) UNIV MINNESOTA.  
PA  
XX Wackelt LP, Mandelbaum RT;  
PI  
XX WPI; 1996-208726/21.  
DR  
XX  
PT Biologically pure culture of atrazine-degrading Pseudomonas - useful to  
PT detoxify atrazine, e.g. in soil, at a wide variety of concns.  
XX  
XX  
PS Example 2; Col 33-36; 34pd; English.  
XX  
CC Novel bacterial strain ADP (ATCC 55464), isolated from atrazine-  
CC contaminated soil, is capable of degrading s-triazine cpds., including  
CC atrazine. In an attempt to identify the strain, the 16S ribosomal RNA  
CC sequence (AAT18760) was compared to that of *Escherichia coli* (AAT18759),  
CC Pseudomonas citrometabolis ATCC 13674 (AAT18761-63), Pseudomonas  
CC aeruginosa (AAT18764), Pseudomonas testosteroni (AAT18765) and  
CC Pseudomonas cepacia (AAT18766). It was concluded that ADP is closely  
CC related to, but distinct from, *P. citrometabolis* and *P. aeruginosa*.  
CC (Updated on 16-Oct-2003 to standardise OS field)  
CC  
XX  
SQ Sequence 1536 BP; 384 A; 355 C; 482 G; 0 T; 314 U; 1 Other;  
Query Match 76.4%; Score 1113.4; DB 2; Length 1536;  
Best Local Similarity 69.9%; Pred. No. 0;  
Matches 1021; Conservative 250; Mismatches 181; Indels 9; Gaps 4;  
QY 1 ATTGAACGCTGGGGCAT-GCTTACATGCAAGTGAACGGAGCAGATGCTTGA 59  
DB 28 AUNGAACGCTGGGGCATGCTTACATGCAAGTGAACGGAGCAGATGCTTGA 84  
QY 60 TTTGGTGGCGATGGCGGCGAGTGAATGCAATCGGAACGTAATCGAAGAGGGGGT 119  
DB 85 -AUGCGACGAGUGCGGAGCGGAGUAAUACAUCCGAAAGUGCTTAAGUAGUGGGGAGU 143  
QY 120 AACGATCGAAAGATGCTATACGGCATATCTCAAGAGGAAAGCAGGGGATGAA 179  
DB 144 AACUACUCGAAAGAGUACUAAUACCGCAUGAGAUUCAAGAUAAAGAGGAGCTTUC 203  
QY 180 AGACCTTGGCTTTTGGAGCGCGCATGCTGATTAAGTGGTGGGTTAAAGGCTTA 239  
DB 204 GGGCTTUGUCUACUAGAGCGGCGUAGGCGAGUUAAGUAGUGUGGUAAGGCUUA 263  
QY 240 CCAAGGCGAGATCAAGTGGTCTGAGAGAGCAGCAGCAGCAGCTGGGAGCTGAGAC 299  
DB 264 CCAAGCCUGAGUCUAGUGGUGUAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGC 323  
QY 300 GGGCCGAGCTCTCAAGCGGAGGAGCAGCAGTGGGAAATTTTGAAGTGGGCGCAGCTGAT 359  
DB 324 GGGCCGAGCTCTCAAGCGGAGGAGCAGCAGTGGGAAATTTTGAAGTGGGCGCAGCTGAT 383  
QY 360 CCAAGCAATGCGCGGTGAGTGAAGAGGCTTTCGAGTGAAGGCTTTCGAGTGAAG 419  
DB 384 CCAAGCAATGCGCGGTGAGTGAAGAGGCTTTCGAGTGAAGGCTTTCGAGTGAAG 443

QY 420 AAAAGGTTACGGTAATATCGTGAACCATGACGGATGACAGAGAAAGACCGGCTTA 479  
DB 444 AAAAGGTTACGGTAATATCGTGAACCATGACGGATGACAGAGAAAGACCGGCTTA 503  
QY 480 CTACGTGCGAGAGCGCGGTAAATGTAAGGTGCAAGGCTTAATGGAATTAATGAGGCG 539  
DB 504 CUAAGGCGAGAGCGCGGTAAATGTAAGGTGCAAGGCTTAATGGAATTAATGAGGCG 563  
QY 540 TAAAGGCGAGAGCGCGGTAAATGTAAGGTGCAAGGCTTAATGGAATTAATGAGGCG 599  
DB 564 UAAAGCGUGCGAGGCGGTAAATGTAAGGTGCAAGGCTTAATGGAATTAATGAGGCG 623  
QY 600 TTGCGTTAAACTAAGAGGCTTAAGGTGCAAGGCTTAATGGAATTAATGAGGCG 659  
DB 624 UGCGCAUUGGAGUCGAGAGGCTTAAGGTGCAAGGCTTAATGGAATTAATGAGGCG 683  
QY 660 GTGAAATGCGTGAAGATGTAAGAAAGCATGATGCGGAAGGCGCTCTGGTTAAAC 719  
DB 684 GUAAGGCGTGAAGATGTAAGAAAGCATGATGCGGAAGGCGCTCTGGTTAAAC 743  
QY 720 TGACGCTCATGCAAGAGCGTGGGAGCAACAGGATTAATGATCCCTGATGTCACGC 779  
DB 744 UGACGCTCATGCAAGAGCGTGGGAGCAACAGGATTAATGATCCCTGATGTCACGC 803  
QY 780 CTTAAAGGATGCAATGATGTTGGGCTTAACTTAAGGTTAAAGGTTAAAGGTTAAAG 839  
DB 804 CTTAAAGGATGCAATGATGTTGGGCTTAACTTAAGGTTAAAGGTTAAAGGTTAAAG 863  
QY 840 AAGTTAGCGCGTGGGAGTACGCTGCAAGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 899  
DB 864 AAGTTAGCGCGTGGGAGTACGCTGCAAGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 923  
QY 900 CGCAAGCGGCGGTAAATGTAAGGTGCAAGGCTTAATGGAATTAATGAGGCG 959  
DB 924 CGCAAGCGGCGGTAAATGTAAGGTGCAAGGCTTAATGGAATTAATGAGGCG 983  
QY 960 TGACGCTCATGCAAGAGCGTGGGAGCAACAGGATTAATGATCCCTGATGTCACGC 1016  
DB 984 UGACGCTCATGCAAGAGCGTGGGAGCAACAGGATTAATGATCCCTGATGTCACGC 1043  
QY 1017 GCTGATGCGCTGCTGACGCTGCTGCTGAGATGTTGGTTAAAGTCCGCAAGCGCG 1076  
DB 1044 GCTGATGCGCTGCTGACGCTGCTGCTGAGATGTTGGTTAAAGTCCGCAAGCGCG 1103  
QY 1077 AACCTTGTATTAATGCTATGTTGGGCGCTTAAATGAGATGCGGCTGACAA 1136  
DB 1104 AACCTTGTATTAATGCTATGTTGGGCGCTTAAATGAGATGCGGCTGACAA 1162  
QY 1137 ACCGAGGAAAGTGGGAGTGAAGTCAAGTCTCAATGAGGCTTAATGAGGCTTAATGAGG 1196  
DB 1163 ACCGAGGAAAGTGGGAGTGAAGTCAAGTCTCAATGAGGCTTAATGAGGCTTAATGAGG 1222  
QY 1197 GTAAATCAATGAGGCTTAATGAGGCTTAATGAGGCTTAATGAGGCTTAATGAGG 1256  
DB 1223 GTAAATCAATGAGGCTTAATGAGGCTTAATGAGGCTTAATGAGGCTTAATGAGG 1282  
QY 1257 GCGCTGTAAGTCCGATTCGAGTCTGCAACTGCACTCGTGAAGTGGAAATGCTTGA 1316  
DB 1283 CCAAGUGUAGUCGAGUCCGAGUCUCAAACUCCAGUCGUGAAGUGGAAUCCUAGUAA 1342  
QY 1317 TCGCGGATCAGATGTCGCGGTGAATGATCCCGGCTTGAAGTGAAGTGAAGTGAAGTGAAG 1376  
DB 1343 TCGCGGATCAGATGTCGCGGTGAATGATCCCGGCTTGAAGTGAAGTGAAGTGAAGTGAAG 1402  
QY 1377 CCAATGGAAGTGGGTTCAACAGAGAGGTAAGTCAATGGAAGAGGCGCTTGGCAGG 1436  
DB 1403 CCAATGGAAGTGGGTTCAACAGAGAGGTAAGTCAATGGAAGAGGCGCTTGGCAGG 1462  
QY 1437 GTGAGATTCATGATGAGGTTG 1457  
DB 1463 GTGAGATTCATGATGAGGTTG 1483

## RESULT 62

AAAF74542

ID AAAF74542 standard; DNA; 1451 BP.

XX AAAF74542;

XX 09-MAY-2001 (first entry)

XX Burkholderia sp. 16S rRNA nucleotide sequence SEQ ID NO:1.

DE Burkholderia; 16S rRNA; terephthalic acid; isophthalic acid; TPA;  
KW p-xylylene; m-xylylene; 4-methylbenzyl alcohol; para-tolualdehyde; polyester;  
KW fibre; film; paint; adhesive; beverage container; ds.

XX Burkholderia sp.

XX US6187569-B1.

XX 13-FEB-2001.

XX 29-JUN-1999; 99US-00342579.

XX 02-JUL-1998; 98US-0091645P.

XX (DUPO ) DU PONT DE NEMOURS &amp; CO E. I.

XX Brumucci MG, Mccutchen CM, Nagarajan V, Thomas SM;

XX WPI; 2001-234334/24.

PT Production of terephthalic acid useful for the production of polyester  
involves bioconversion of p-xylylene comprises a single terephthalic acid  
producing microorganism.

XX Claim 4; Col 19-22; 21pp; English.

XX The present invention describes a method for the production of a  
CC terephthalic acid (TPA) comprising contacting an isolated TPA producing  
CC Proteobacteria with an aromatic organic substrate selected from p-xylylene,  
CC 4-methylbenzyl alcohol and para-tolualdehyde and recovering TPA. The  
CC method can be used for the production of polyesters which are required  
CC for fibers, films, paints, adhesives and beverage containers. The process  
CC is environmentally friendly, safe, economical and produces compounds of  
CC commercial interest. The method has broad applicability for production of  
CC terephthalic acid and isophthalic acid. The present sequence represents  
CC the specifically claimed 16S rRNA nucleotide sequence from the  
CC Burkholderia sp. TPA producing microorganism

XX Sequence 1451 BP; 379 A; 339 C; 441 G; 292 T; 0 U; 0 Other;

XX Query Match 75.8%; Score 1104.8; DB 4; Length 1451;  
XX Best Local Similarity 87.0%; Pred. No. 0;  
XX Matches 1250; Conservative 0; Mismatches 177; Indels 9; Gaps 3;

QY	11	GGCGGCATGCTTTACATGCAAGTCGAACGCGACAGCGATGCTTGATCTGTGGCGA	70
DB	11	GGCGGCATGCTTTACATGCAATTCMAACGCGACAGCGATGCTTGATCTGTGGCGA	70
QY	71	GTGGCGACGGGTGAGTAATGATCGGAACTATCCAGAAAGCGGGGTAAACGATCCGA	130
DB	71	TTGGCGAAGGGGTGATTAATCATCGGAATGTAACCTTGAGTGGGGGATACCTCGGC	130
QY	131	AGATGTGCTAATACCGCATATATCTTAAGGAGAAAGCGGGATCGAAAGACTTGGCG	190
DB	131	AGCCGATTAATACCGCATACGCTCTGAGAGAGAAAGCGGGGACCTTGGGCTTCGCGC	190
QY	191	TTTGGAGCGCGCGATGCTGATTAAGCTAGTGTGGGTAAAGGCTTACCAAGCGAG	250
DB	191	TACAAAGACGCGATGCTAATTAATCTATTTGGTGGGTAAAGCTTACCAAGCGAG	250
QY	251	ATCAGTAGTGTGCTGAGAGACGACCGCACTGGGACTGAGACAGCGGCCAGATCTC	310

DB	251	ATCAGTAGTGTGCTGAGAGACGACCGCACTGGGACTGAGACAGCGGCCAGATCTC	310
QY	311	CTACGGGAGCGACAGTGGGGAAATTTTGGCAATGGGCGCAAGCTTGAATCAGCAATGCC	370
DB	311	CTACGGGAGCGACAGTGGGGAAATTTTGGCAATGGGCGCAAGCTTGAATCAGCAATGCC	370
QY	371	GGCGTGAAGAAAGGCGCTTGGGGTGTAAAGCTCTTCAAGTGAAGAAAGGTTACG	430
DB	371	GGCGTGTGAAGAAAGGCGCTTGGGGTGTAAAGCACTTTTGTCCGAAAGAAATCCTCTCG	430
QY	431	GTAATATATCGTGAACCATGACGATATCGACAGAAAGACCGGCTTAATCACTGTCAG	490
DB	431	GTTAATATCTCCGGGGGATGACGCTACCGGAAATAAGACCGGCTTAATCACTGTCAG	489
QY	491	CAGCGCGGTATATCGTAGGGTGCAGCGTTAATCGGAATTACGCGCGTAAAGGTCCG	550
DB	490	CAGCGCGGTATATCTTAGGGTGCAGCGTTAATCGGAATTACGCGCGTAAAGCGTCCG	549
QY	551	CAGCGCGCTTGTATGTCAGATGTAATCCCGGGCTTAACTGGGAAATTCGTTTGA	610
DB	550	CAGCGCGTGTGTGAAGCGAGATGTAATCCCGGGCTTAACTGGGAAATTCGTTTGTG	609
QY	611	ACTCAAGACTAGAGTGTGCAGAGAGGAGGTGAATTCATGTGTAGCAGTGAATGCGT	670
DB	610	ACTCAAGGCTAGAGTGTGCAGAGAGGAGGTGAATTCATGTGTAGCAGTGAATGCGT	669
QY	671	AGAGATATGAAAGACATCGATGCGAAAGCGCTCTTGCTTAACTGACCTCATG	730
DB	670	AGAGATGTGAGAAATCCGATGCGAAAGCGCTCTTGCTTAACTGACCTCATG	729
QY	731	CACGAAAGCGTGGGAGCAAAACAGATTGAATACCTGTGTATGTCACGCGCTTAAAGATG	790
DB	730	CACGAAAGCGTGGGAGCAAAACAGATTGAATACCTGTGTATGTCACGCGCTTAAAGATG	789
QY	791	TCAACTAGTGTGGGCTTATTAAGGCTGTGTAACGAAGCTAACGCGTGAAGTTACCGC	850
DB	790	TCAACTAGTGTGGGCTTATTAAGGCTGTGTAACGAAGCTAACGCGTGAAGTTACCGC	849
QY	851	CTGGGAGTACGCTGCGCAAGTTTAACTCAAGAAATTGACGGGACCGCGCAACAGGG	910
DB	850	CTGGGAGTACGCTGCGCAAGTTTAACTCAAGAAATTGACGGGACCGCGCAACAGGG	909
QY	911	TGATTTATGCGATTAATTCGATGCAACGCGAAACCTTACCTGATGATGTAGC	970
DB	910	TGATTTATGCGATTAATTCGATGCAACGCGAAACCTTACCTGATGATGTAGC	969
QY	971	GAATTTCTAGATAGTATGATGCT--TCGGGACGCTTAACACAGGTGCTGATGCT	1027
DB	970	GAATTTCTAGATAGTATGATGCT--TCGGGACGCTTAACACAGGTGCTGATGCT	1029
QY	1028	GTCGTACGCTGCTGCTGATGATGCTTGAAGTCCCGCAACGAGCGCAACCTTGTCA	1087
DB	1030	GTCGTACGCTGCTGCTGATGATGCTTGAAGTCCCGCAACGAGCGCAACCTTGTCC	1089
QY	1088	TTAATTCGATCATTTGTTGGGCACTTATATGACCTGCGGGTGAACAAACCGAGGAG	1147
DB	1090	TTAGTTGC-----TACGGAAGACATCTTAAGGACCTGCGGGTGAACAAACCGAGGAG	1144
QY	1148	GTGGGAGTACGTCAGAGCTCATGAGCTTATGGGTAGGGCTTACACGATTAACAATG	1207
DB	1145	GTGGGAGTACGTCAGAGCTCATGAGCTTATGGGTAGGGCTTACACGATTAACAATG	1204
QY	1208	GGCGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATTCAGAAAGCGCGTGTATG	1267
DB	1205	GTGGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATTCAGAAAGCGCGTGTATG	1264
QY	1268	CCGATTCGGAATTCGCACTGACCTCCGTGAAGTCCGGAATGAGTAAATGCGCGATAG	1327
DB	1265	CCGATTCGGAATTCGCACTGACCTCCGTGAAGTCCGGAATGAGTAAATGCGCGATAG	1324
QY	1328	CATGTCGCGGTGAATACGTTCCCGGCTTGTACACACGCGCGCTGACACCATGGAGTG	1387
DB	1325	CATGTCGCGGTGAATACGTTCCCGGCTTGTACACACGCGCGCTGACACCATGGAGTG	1384

QY 1388 GGTTCACCGAAGCAGCTGATCTTAACCTTAAGAGGCGCTTGCCACGGTGAGAT 1443  
DB 1385 GGTTCACCGAAGCAGCTGATCTTAACCTTAAGAGGCGCTTGCCACGGTGAGAT 1440

## RESULT 63

ABX10819  
ID ABX10819 standard; DNA; 1451 BP.

ABX10819;

17-APR-2003 (first entry)

DE Burkholderia sp 16S ribosomal RNA gene.

XX Ribosomal RNA; rRNA; 16S; terephthalic acid; TA; proteobacteria;  
KM TA biosynthetic pathway; isophthalic acid; 4-carboxybenzyl alcohol;  
KM 4-carboxybenzaldehyde; fibre; film; paint; adhesive; beverage container;  
KM gene; ds.

OS Burkholderia sp.

PN US6461840-B1.

PD 08-OCT-2002.

PF 17-JUL-2000; 2000US-00617854.

PR 02-JUL-1998; 98US-0091645P.

PR 28-JUN-1999; 99US-00342579.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Bramucci MG, Mccutchen CM, Nagarajan V, Thomas SM;

XX WPI; 2003-208774/20.

XX Terephthalic acid producing proteobacteria isolated by culturing sample  
PT suspected of containing the bacteria in medium having aromatic organic  
PT substrate, adding xylene and isolating bacteria producing terephthalic  
PT acid.

PS Claim 4; Col 19-20; 21pp; English.

XX The invention describes a pure isolate of a terephthalic acid (TA)  
CC producing proteobacteria, comprising genes encoding all enzymes of the TA  
CC biosynthetic pathway. The proteobacteria are useful for producing TA and  
CC isophthalic acid, and various intermediates in the synthesis of TA such  
CC as 4-carboxybenzyl alcohol and 4-carboxybenzaldehyde. The TA and  
CC isophthalic acid have utility in the production of polyesters needed in  
CC fibres, films, paints, adhesives and beverage containers. This sequence  
CC represents the Burkholderia sp 16S ribosomal RNA gene amplification of  
CC which was used to identify Burkholderia and Pseudomonas isolates

XX Sequence 1451 BP; 379 A; 339 C; 441 G; 292 T; 0 U; 0 Other;

XX Query Match 75.8%; Score 1104.8; DB 10; Length 1451;

XX Best Local Similarity 87.0%; Pred. No. 0;

XX Matches 1250; Conservative 0; Mismatches 177; Indels 9; Gaps 3;

QY 11 GGGGGCATGCTTTACACATGCAAGTCCGACGCGACGATGCTTGCATCTGTTGGCGA 70  
DB 11 GGGGGCATGCTTTACACATGCAAGTCCGACGCGACGATGCTTGCATCTGTTGGCGA 70  
QY 71 GTGGCGGACGGGTGAGTATGATCGGAACGTATCCGAAAGAGGGGGTAAACGATCGAA 130  
DB 71 GTGGCGGACGGGTGAGTATGATCGGAACGTATCCGAAAGAGGGGGTAAACGATCGAA 130  
QY 131 AGATGTGCTAATACCGCATTAATCTTAAGAGGAAACGAGGATCGAAAGACCTTGGCG 190  
DB 131 AGCGGATTAATACCGCATTAATCTTAAGAGGAAACGAGGATCGAAAGACCTTGGCG 190

QY 131 AGCGGATTAATACCGCATTAATCTTAAGAGGAAACGAGGATCGAAAGACCTTGGCG 190

QY 191 TTTTGAGCGGCGCATGTCTGATGCTAGTTAGTGTGGGGTAAAGGCTTCAAGCGGACG 250  
DB 191 TACAAAGACGCGCATGTCTGATGCTAGTTAGTGTGGGGTAAAGGCTTCAAGCGGACG 250  
QY 251 ATCAGTAGTGTGTCTGAGAGGACGACGACGACCTGAGACACGCGCCAGATCTC 310  
DB 251 ATCTGTACTGTGTCTGAGAGGACGACGACGACCTGAGACACGCGCCAGATCTC 310  
QY 311 CTACGGGAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 370  
DB 311 CTACGGGAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 370  
QY 371 GCGTAGTAGAGAGAGGCTTGGGTTGTAAAGCTTTTCACTGAGAGAGAGAGGTTAGC 430  
DB 371 GCGTAGTAGAGAGAGGCTTGGGTTGTAAAGCTTTTCACTGAGAGAGAGAGGTTAGC 430  
QY 431 GTAAATATCTGACCCATGACGATGATGCAAGAGAGACGCGCTTAACTACGTCGAC 490  
DB 431 GTAAATATCTGACCCATGACGATGATGCAAGAGAGACGCGCTTAACTACGTCGAC 490  
QY 491 CAGCGCGGTAATATCTGAGGTCGACGCTTAACTGAGAGAGAGAGAGAGAGAGAGAG 550  
DB 491 CAGCGCGGTAATATCTGAGGTCGACGCTTAACTGAGAGAGAGAGAGAGAGAGAGAG 550  
QY 551 CAGCGCGGCTTGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610  
DB 551 CAGCGCGGCTTGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610  
QY 611 ACTACAAAGCTAG 670  
DB 611 ACTACAAAGCTAG 670  
QY 671 AGAGATATGAG 730  
DB 671 AGAGATATGAG 730  
QY 731 CACGAAGCGTGGGAGCAAAACAGATTAATCTGAGTATCTGAGTATCTGAGTATCTG 790  
DB 731 CACGAAGCGTGGGAGCAAAACAGATTAATCTGAGTATCTGAGTATCTGAGTATCTG 790  
QY 791 TCACTAGTGTGGGCTTATTAATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAG 850  
DB 791 TCACTAGTGTGGGCTTATTAATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAG 850  
QY 851 CTGGGAGTACGCTGCGCAAGATTAATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 910  
DB 851 CTGGGAGTACGCTGCGCAAGATTAATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 910  
QY 911 TGATATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 970  
DB 911 TGATATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 970  
QY 971 GAATTTTCTAGAGATGATTAATGATGATGATGATGATGATGATGATGATGATGATG 1027  
DB 971 GAATTTTCTAGAGATGATTAATGATGATGATGATGATGATGATGATGATGATGATG 1027  
QY 1028 GTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1087  
DB 1028 GTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1087  
QY 1088 TTAATGCAATCAATTTGTTGGGCACTTAAATGAGATGAGATGAGATGAGATGAG 1147  
DB 1088 TTAATGCAATCAATTTGTTGGGCACTTAAATGAGATGAGATGAGATGAGATGAG 1147  
QY 1148 GTGGGAGTACGCTCAAGTCTTCAATGAGCTTAAATGAGATGAGATGAGATGAGAT 1207  
DB 1148 GTGGGAGTACGCTCAAGTCTTCAATGAGCTTAAATGAGATGAGATGAGATGAGAT 1207  
QY 1208 GCGGCTACAGAGGTTGCCAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267  
DB 1208 GCGGCTACAGAGGTTGCCAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267  
QY 1268 CCGGATCGAGAGTGTGCACTGCACTCCGTGAAGTCCGATTAATCCGAGATCGAG 1327  
DB 1268 CCGGATCGAGAGTGTGCACTGCACTCCGTGAAGTCCGATTAATCCGAGATCGAG 1327

Db	1265	CCGAGTCCAGCTCGCAACTGCACTGCGCTGTAAGCTGGAATGCTTAGTGAATGCGGATCAG	132/24
Oy	1328	CATGTCGCGGGAATACGCTTCCGGGCTTTTGACACACCGCCCGTACACCATGGAGTG	138/7
Db	1325	CATGTCGCGGGAATACGCTTCCGGGCTTTTGACACACCGCCCGTACACCATGGAGTG	138/4
Oy	1388	GATTTCACACGAGACGAGTACTTAAACCGTAAAGAGGGCGCTTGGCCACAGGTGAGAT	144/3
Db	1385	GATTTCACAGAAAGTAGGTAGCTTAACCGTAAAGAGGGCGCTTACACGCGCAGAT	144/0
RESULT	64		
ID	ADH74483		
XX	ADH74483	standard; DNA; 1451 BP.	
AC	ADH74483;		
XX	25-MAR-2004	(first entry)	
DT			
XX	TPA	producing microorganism isolate IR3 16s rRNA gene.	
DE			
XX	Terephthalic acid; TPA	producing microorganism; p-xylene; p-toluic acid;	
KM	isophthalic acid; IPA;	polyester; fibre; film; paint; adhesive;	
KM	isophthalic acid; IPA;	polyester; fibre; film; paint; adhesive;	
KW	beverage container; condensation reaction;	diamine; polyamide; 16s rRNA;	
XX	IR3; ds.		
XX			
OS	Burkholderia sp.		
XX			
PN	US2003170836-A1.		
XX			
PD	11-SEP-2003.		
XX			
PF	15-AUG-2002; 2002US-00219549.		
XX			
PR	02-JUL-1998; 98US-0091645P.		
PR	29-JUN-1999; 99US-00342579.		
PR	17-JUL-2000; 2000US-00617854.		
XX			
PA	(BRAM/) BRAMUCCI M G.		
PA	(MCCU/) MCCUTCHEN C M.		
PA	(NAGA/) NAGARAJAN V.		
PA	(THOM/) THOMAS S W.		
PI	Bramucci MG, Mccutchen CM, Nagaraajan V, Thomas SM;		
XX			
DR	WPI; 2004-096737/10.		
XX			
PT	New terephthalic acid-producing microorganism e.g. for producing		
PT	terephthalic acid for production of polyesters commercially required in		
PT	large quantities for fibers, films, paints, adhesives and beverage		
PT	containers.		
XX			
PS	Claim 6; SEQ ID NO 1; 24pp; English.		
XX			
XX			
CC	The invention relates to a terephthalic acid (TPA) producing		
CC	microorganism isolated by a process comprising culturing a sample		
CC	suspected of containing a TPA producing microorganism in a suitable		
CC	growth medium containing at least one aromatic organic substrate selected		
CC	from p-xylene, p-toluic acid and TPA, selecting those microorganisms that		
CC	are able to use every substrate selected individually from p-xylene, p-		
CC	toluic acid and TPA as a sole carbon source, contacting the		
CC	microorganisms selected in the above step with p-xylene to form a		
CC	reaction medium and monitoring the reaction medium over time for the		
CC	presence of TPA, where the production of TPA indicates the presence of a		
CC	TPA producing microorganism. The invention also relates to a method of		
CC	producing isophthalic acid (IPA) by contacting an isolated TPA producing		
CC	microorganism with an aromatic organic substrate and at least one		
CC	suitable induction compound, where isophthalic acid accumulates, and		
CC	optionally recovering the isophthalic acid. The microorganism is useful		
CC	for producing terephthalic acid for the production of polyesters that are		
CC	commercially required in large quantities for fibres, films, paints,		
CC	adhesives and beverage containers. The isophthalic acid is useful in		

CC	condensation reactions with diamines to form polyamides. This sequence
CC	represents the 16S rRNA gene of the TPA producing microorganism
CC	Burkholderia isolate IR3.
XX	
SQ	Sequence 1451 BP; 379 A; 339 C; 441 G; 292 T; 0 U; 0 Other;
Query Match	75.8%; Score 1104.8; DB 12; Length 1451;
Best Local Similarity	87.0%; Pred. No. 0;
Matches 1250;	Conservative 0; Mismatches 177; Indels 9; Gaps 3
Dy	11 GGCGGCATGCTTTACATGCAAGTCGAACGGCAGACGATGCTTGCACTGTGGCGA 70
Dd	11 GGCGGCATGCTTTACATGCAATTCAAACGGCAGACGAGTCTTGCACTGTGGCGA 70
Dy	71 GTGGCGACGGGTGATATGATCATTCGGAACGTAATCCAGAAGGGGGTTAACGATCGAA 130
Dd	71 TTGGGGAACGGGTGATTATATCATTCGGAATATCACTTATGTGGGGATACCTCGCAAA 130
Dy	131 AGATGTGTATAACCGCATATCTCTAGAGAGAAAGCAGGGGATCGAAAAGACTTGTGGC 190
Dd	131 AGCGGATTAATACCGCATACGCTCTTGAGAGAAAGCGGGGACCTTCGGGCTTCGCC 190
Dy	191 TTTTGGACGGGCCGATGTCTGATTAGCTAGTGTGGGTAAAGGCTTACCAGGCGACG 250
Dd	191 TACAAAACACCCGATGTCAATTATCTATTGTGGGTAAAGCTCACCAAAGCGACA 250
Dy	251 ATCAGTAGTGGTCTTGAGAGACGACCAACCACTGGGACTGAGACACGGGCCAGATCTC 310
Dd	251 ATCTGTACTGGTCTGAGAGGCAACCAACCACTGGGACTGAAACACGGGCCAAATCTC 310
Dy	311 CTACCGGAGGCGACGATGGGGGATTTTGGACAAATGGGGCGAAGCTGATCCAGATGCG 370
Dd	311 CTACCGGAGGCGACGATGGGGGATTTTGGACATGGGGCGAAGCTGATCCACATGCG 370
Dy	371 GCGTAGTGAAGAAAGGCTCTGGGTGTAAAGCTCTTCAGTCGAGAAAGAAAGTTACG 430
Dd	371 GCGTGTGTAAAAAAGGCTCTGGGTGTAAAGCACTTTGTCCGAAAGAAATCTCTGG 430
Dy	431 GTAAATTAATCTGACCCATGACGATATCGACAGAGAAACACCGGCTAATCTACGTGCCAG 490
Dd	431 GTTAATTAATCTCGAGGGGATGACGGTATCCGGAATAAGACACCGGCTAATCTCTGCCA- 489
Dy	491 CAGCCGCGGTAAATACGTAGGGTGCNAAGCTTAAATGGGAGTGAAGAGGTGG 550
Dd	490 CAGCCGCGGTAAATCTAGGGTGCNAAGCTTAAATGGGAGTGAAGAGGTGG 549
Dy	551 CAGGGGGGCTTTAATGTGATGTGAATCCCCGGGCTTAACTGGGAATTCGTTTGA 610
Dd	550 CAGGGGGGTTTTTAAGACGATGTGAATCCCCGGGCTTAACTGGGAATTCGATTGCG 609
Dy	611 ACTACAAAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTGACATGAATGCGT 670
Dd	610 ACTGCAAGCTAGAGTATGGCAGAGGGGGGTGAATTCACGTGTAGCAGTGAAGTGGT 669
Dy	671 AAGATATGAAGAAATGATGTGGGCAAGGACGCTCTGGGGTTAACATGACGCTCATG 720
Dd	670 AAGATATGAAGAAATGATGTGGGCAAGGACGCTCTGGGGTTAACATGACGCTCATG 729
Dy	731 CACGAAAGGTGGGGGAGCAACAGGATTGATACCTGTGTAGTCCAAGGCTTAAAGATG 790
Dd	730 CACGAAAGGTGGGGGAGCAACAGGATTGATACCTGTGTAGTCCAAGGCTTAAAGATG 789
Dy	791 TCACACTAGTTGTGTGGGCTTATTAAGCTTGTGTAACGACCTAAACGCTGAAGTTGACCG 850
Dd	790 TCACACTAGTTGTGTGGGATTCATTTCTTAGTAACAACTAAACGCTGAAGTTGACCG 849
Dy	851 CTGGGGAGTACGCTGCGCAAGTTAAACCTCAAGGAATTGACGGGGACCCGCAAGCGG 910
Dd	850 CTGGGGAGTACGCTGCGCAAGTTAAACCTCAAGGAATTGACGGGGACCCGCAAGCGG 909
Dy	911 TGGAATTAATGATTAATTCGATGCAACGCGAAGAACTTACCTACCTTGAATGTAGC 970
Dd	910 TGGAATTAATGATTAATTCGATGCAACGCGAAGAACTTACCTACCTTGAATGTAGC 969



QY 971 GAATTTCTAGAGATAGATTAGTCT---TCGGGAAACGCTAACACAGCTCTGCATGGCT 1027  
 CC 970 GAATCTCTGTAGAGAGGTGAGAGTCTGAAAGGAAACCGTAAACACAGGTCTGCATGGCT 1029  
 Db 1028 GTCTGACGCTCGGTGTGATGATGTTGGGTTAAGTCCCGCAACGACGCAACCTTGTCA 1087  
 QY 1030 GTCTGACGCTCGGTGTGATGATGTTGGGTTAAGTCCCGCAACGACGCAACCTTGTCC 1089  
 Db 1088 TTAAATTCGATCATTTTGTGGGCACTTTAATGAGACTGCGGTGACAAACCGAGAGAA 1147  
 Db 1090 TTAGTTG-----TACGCAAGACACTCTAAAGAGACTGCGGTGACAAACCGAGAGAA 1144  
 QY 1148 GTGGGATGACGCTAGCTCTCATGAGCCCTTAATGGGTTCACAGCTAATTAAGAAG 1207  
 Db 1145 GTGGGATGACGCTAGCTCTCATGAGCCCTTAATGGGTTCACAGCTAATTAAGAAG 1204  
 QY 1208 GCGCGTACAGAGGGTGGCCAAACCGGAGGGAGGCTTAATCTCAGAAAGCGCTGTAGT 1267  
 Db 1205 GTCTGACAGAGGGTGGCCAAACCGGAGGGAGGCTTAATCTCAGAAAGCGCTGTAGT 1264  
 QY 1268 CCGGATCGAGTCTGCACTCGACTCGTGAAGTCGGAATCGCTAGTAATCGCGATCAG 1327  
 Db 1265 CCGGATCGAGTCTGCACTCGACTCGTGAAGTCGGAATCGCTAGTAATCGCGATCAG 1324  
 QY 1328 CATGTGCGCGTGAATACGTTCCCGGCTTTGTACACACCGCCCGTCAACATGGAGTG 1387  
 Db 1325 CATGTGCGCGTGAATACGTTCCCGGCTTTGTACACACCGCCCGTCAACATGGAGTG 1384  
 QY 1388 GGTTCACAGAAAGAGTGTCTTAACCGTAAAGAGGGCGCTTACCGACGGTAGAT 1443  
 Db 1385 GGTTCACAGAAAGTGTGTACCTTAACCGTAAAGAGGGCGCTTACCGACGGTAGAT 1440

RESULT 65  
 AAX16349  
 ID AAX16349 standard; DNA; 1455 BP.  
 XX AAX16349;  
 AC 23-APR-1999 (first entry)  
 XX 16S ribosomal DNA #1 from a bacterium strain A-1.  
 DE 16S ribosomal DNA; gram-negative aerobic filamentous bacterium;  
 KM strain A-1; intracellular granule; poly-beta-hydroxybutyrate;  
 KM chlorinated aliphatic compound; biological degradation; ss.  
 XX Unidentified.  
 OS US5874291-A.  
 PN 23-FEB-1999.  
 PD 30-APR-1996; 96US-00642229.  
 XX 20-MAY-1994; 94US-00246865.  
 PR 12-FEB-1996; 96US-00599867.  
 XX (UNITW ) UNIV WASHINGTON.  
 PA Herwig RP, Strand SE, Bielefeldt AR, Stenel HD;  
 PI WPI; 1999-179986/15.  
 DR Biodegradation of chlorinated aliphatic compounds - with Gram-negative  
 PT aerobic filamentous bacterium.  
 XX Claim 3; Col 23-26; 35pp; English.  
 PS The present sequence represents the 16S ribosomal DNA from an unknown  
 CC (possible new species) gram-negative aerobic filamentous bacterium,  
 CC designated strain A-1. The present invention also describes: (1) a

CC bioremediation method comprising contacting an environmental sample  
 CC contaminated with chlorinated aliphatic compounds (CACs) with a CAC-  
 CC metabolising aerobic Gram-negative bacterium which: (a) has non-sheathed  
 CC rod-shaped cells that are filamentous in appearance and have a length of  
 CC 20-200 micro m; (b) accumulates poly-beta-hydroxybutyrate in  
 CC intracellular granules under nitrogen-limiting conditions; (c) is capable  
 CC of growing in the presence of 0.2% phenol or 1% saturated toluene vapour;  
 CC and (d) does not grow in carbohydrates; (2) a bioremediation method  
 CC comprising selecting a soil or water sample contaminated with CACs,  
 CC contacting the sample with a stream of air into which the CACs are  
 CC transferred, and contacting the CAC-bearing air with a bacterium as  
 CC above. The methods are to degrade CACs such as trichloroethylene (TCE)  
 CC and dichloroethylene (DCE) in soil, water or air. The bacterium is more  
 CC suitable for bioremediation than methanotrophic bacteria due to its  
 CC ability to tolerate CAC metabolites, its tolerance for relatively high  
 CC concentrations of substrate and its having a higher rate of endogenous  
 CC CAC conversion, and can tolerate higher TCE concentrations than other  
 CC phenol-degrading microorganisms

SQ Sequence 1455 BP; 365 A; 344 C; 464 G; 279 T; 0 U; 3 Other;

Query Match 75.6%; Score 1101.8; DB 2; Length 1455;  
 Best Local Similarity 86.9%; Pred. No. 0;  
 Matches 1270; Conservative 1; Mismatches 179; Indels 12; Gaps 5;

QY 1 ATTTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGGGACACGGATGCTTGAT 60  
 Db 1 ATTTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGGGACACGGATGCTTGAT 58  
 QY 61 CTGTGTGCGAGTGTGGCGGACGCGGTGATATGATCGGAACGTAATCAGAAAGGGGGGTA 120  
 Db 59 CTGTGTGCGAGTGTGGCGGACGCGGTGATATGATCGGAACGTAATCAGAAAGGGGGGTA 118  
 QY 121 ACCGATCGAAAGATGTCTAATACCGCATATCTTAAAGAGAAAGCGGGGATCGAAA 180  
 Db 119 ACTTACTCGAAAGATGTCTAATACCGCATATCTTAAAGAGAAAGCGGGGATCGCAA 178  
 QY 181 GACCTTGGCCTTTTGGAGCGCGCGATGCTGATTAAGTGTGAGGTTAAAGGCTTAC 240  
 Db 179 GACCTTGGCCTTACGAGCGCGCTGCTGAGTATTAAGTGTGAGGTTAAAGGCTTAC 238  
 QY 241 CAAAGCGACGATCAGTATGTTGTTGAGAGACGACAG-CCACATGCGGACTGAGACAC 299  
 Db 239 CAAAGCGACGATCAGTATGTTGTTGAGAGACGACAG-CCACATGCGGACTGAGACAC 298  
 QY 300 GGCCCAAGACTCTTACGCGGAGGACGACAGTGGGAAATTTTGGACAAATGGGGCAAGCTGAT 359  
 Db 299 GGCCCAAGACTCTTACGCGGAGGACGACAGTGGGAAATTTTGGACAAATGGGGCAAGCTGAT 358  
 QY 360 CCAGCAATGCGCGGT-GAGTGAAGAAAGGCTTGGGTTTAAAGCTCTTCAAGTCAGAA 418  
 Db 359 CCAGCAATGCGCGGTGAGTGAAGAAAGGCTTGGGTTTAAAGCTCTTCAAGTCAGAA 418  
 QY 419 GAAAAGGTTACCGTAAATATATCGTGAACCCATGACCGTATCGACAGAAAGACCGGCTTA 478  
 Db 419 GAAAAGGCTCTCTTAATATCAGAGACGATACCGTATCGTAAAGAAATTAACCGGCTTA 478  
 QY 479 ACTTACGTCACAGACCGCGGTAAATCGTAAAGGTTCAACCGTAAATTAACCGGCTTA 538  
 Db 479 ACTTACGTCACAGACCGCGGTAAATCGTAAAGGTTCAACCGTAAATTAACCGGCTTA 538  
 QY 539 GTTAAAGGTTGCGCAAGCGGCTTGTAAAGTCAATGTAATCCCGGGCTTAACTGAGGA 598  
 Db 539 GTTAAAGGTTGCGCAAGCGGCTTGTAAAGTCAATGTAATCCCGGGCTTAACTGAGGA 598  
 QY 599 ATTGGCTTTGAATACTAACAGCTAGAGTGTGACAGAGGGAGGTGAATTCATGTGTAC 658  
 Db 599 ACGGCTTTGTGATCTGCAAGAGTGTGAGTGTGACAGAGGGAGGTGAATTCATGTGTAC 658  
 QY 659 AGTGAATGCGTGAAGATATGGAAGAACTGATGCGGAAGGACGCTCTGTGGTTAA 718  
 Db 659 AGTGAATGCGTGAAGATATGCGAGGAAACCGATGCGGAAGGCAATCCCTGTGGCTTGA 718



```
QY 719 CTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGGTAGTCCACG 778
DB 719 CTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGGTAGTCCACG 778
QY 779 CCCCTAAACGATGTCATCACTAGTGTGGGCCCTTATTAGGCTTGGTAAGAACTAACCGGT 838
DB 779 CCCCTAAACGATGTCATCACTAGTGTGGGCCCTTATTAGGCTTGGTAAGAACTAACCGGT 838
QY 839 GAAAGTTAGCCGCTGGGAGATACCGTGCAGATTAATAAAGTCAAGAAATTTGACGGGAGC 898
DB 839 GAAAGTTAGCCGCTGGGAGATACCGTGCAGATTAATAAAGTCAAGAAATTTGACGGGAGC 898
QY 899 CCGCAGACGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAAACTTACTTACC 958
DB 899 CCGCAGACGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAAACTTACTTACC 958
QY 959 TTGACATGATGACCAATTTCTAGATGATTAATGCTCT--TGGGAAACGTAAACAGAG 1015
DB 959 TTGACATGATGACCAATTTCTAGATGATTAATGCTCTCTGCAACAGAG 1018
QY 1016 TGTGATGATGATGCTGCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1075
DB 1019 TGTGATGATGATGCTGCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1078
QY 1076 CAACCTTGTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135
DB 1079 CAACCTTGTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133
QY 1136 AACCGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1195
DB 1134 AACCGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
QY 1196 CGTAAATCAATGAGCGGTACAGAGGATGACCAACCGGAGGAGGATTAATCTCAGAA 1255
DB 1194 CGTAAATCAATGAGCGGTACAGAGGATGACCAACCGGAGGAGGATTAATCTCAGAA 1253
QY 1256 GCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1315
DB 1254 GCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
QY 1316 ATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375
DB 1314 ATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373
QY 1376 AACATGAGAGTGGGTTTACCAAGAGAGATTAATCCGTAAGAGAGGAGGCTTGCAC 1435
DB 1374 AACATGAGAGTGGGTTTACCAAGAGAGATTAATCCGTAAGAGAGGAGGCTTGCAC 1433
QY 1436 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
DB 1434 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455
```

RESULT 66  
ADG4143/c  
ID ADG4143 standard; DNA; 1453 BP.

AC ADG4143;  
XX 26-FEB-2004 (first entry)  
XX Unknown organism strain RA2 16S rDNA.  
KM de; rDNA; bacterial strain identification; wastewater bioreactor; 16S.  
XX Unidentified.  
OS US2003207321-A1.  
XX PN 06-NOV-2003.  
XX PD 17-JUN-2003; 2003US-00464709.  
XX PF  
XX

```
PR 16-DEC-1999; 99US-0171140P.  
PR 13-DEC-2000; 2000US-00735567.  
XX  
PA (BRAM/) BRAMUCCI M. G.  
PA (CHEN/) CHEN M. W.  
PA (ALBE/) ALBERTSON-KANE H. M.  
PA (NAGA/) NAGARAJAN V.  
XX  
PI Bramucci MG, Chen MW, Albertson-Kane HM, Nagarajan V;  
XX  
DR WPI; 2003-901056/82.  
XX  
PT New isolated 16S rDNA sequence indicative of the presence of an activated  
PT sludge bacterial strain, useful for rapidly identifying and monitoring  
PT bacterial strains in samples taken from a wastewater bioreactor.  
XX  
PS Claim 1; SEQ ID NO 4; 24bp; English.  
XX  
CC The invention relates to an isolated 16S rDNA sequence. The 16S rDNA  
CC sequences and methods are useful in rapidly identifying and monitoring  
CC the corresponding bacterial strains in samples taken from a wastewater  
CC bioreactor or from any environment which might contain the specified  
CC bacteria. The present sequence represents the 16S rDNA from an unknown  
CC organism.  
XX  
SQ Sequence 1453 BP; 279 A; 469 C; 344 G; 361 T; 0 U; 0 Other;  
XX  
Query Match 75.5%; Score 1100; DB 10; Length 1453;  
Best Local Similarity 86.1%; Pred. No. 0;  
Matches 1256; Conservative 0; Mismatches 190; Indels 12; Gaps 3;  
QY 1 ATGACGCTGCGGCAATGCTTTACATGCAATGCAACGGCGACGATGCTTGAT 60  
DB 1449 ATTGAACGCTGCGGCAATGCTTTACATGCAATGCAACGGCGACGATGCTTGAT 1394  
QY 61 CTGATGCGATGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 1393 CTGATGCGATGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334  
QY 121 ACGCATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
DB 1333 GCCCGGAGAAAGCGGATTAATACCGCATGATGATGATGATGATGATGATGATGATGAT 1274  
QY 181 GACCTTGCGCTTTTGGAGCGGCGCATGCTGATGATGATGATGATGATGATGATGATGAT 240  
DB 1273 GGCCTCAACGCTTTGGAGCGGCGCATGCTGATGATGATGATGATGATGATGATGATGAT 1214  
QY 241 CAAGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 1213 CAAGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1154  
QY 301 GCCGAGACTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
DB 1153 GCCGAGACTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1094  
QY 361 CAGCAATGCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTTTCACTGAGAGAA 420  
DB 1093 CAGCAATGCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTTTCACTGAGAGAA 4034  
QY 421 AAGGTTACGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
DB 1033 AAGGTTACGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 974  
QY 481 TACGTGACAGAGCGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
DB 973 TACGTGACAGAGCGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914  
QY 541 AAAGGTGCGAGAGCGGCTTTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 600  
DB 913 AAAGGTGCGAGAGCGGCTTTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 854  
QY 601 TGCCTTGAATCAAGCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
```

Db 853 TGCCTTGTGACTGCAAGGCTGAGTGGCGAGAGGGGAGTGGAAATTCGCGGTGTACAG 794  
 Qy 661 TGAATATGCTAGATATAGAAACATCGATGGCGAGCCTCTGGGTTAACT 720  
 Db 793 TGAATATGCTAGATATAGAAACATCGATGGCGAGCCTCTGGGTTAACT 734  
 Qy 721 GACGCTCATGACGAAAGCGTGGGGGCAAAACAGATTGATCCCTGGAGTCCAGCC 780  
 Db 733 GACGCTCATGACGAAAGCGTGGGGGCAAAACAGATTGATCCCTGGAGTCCAGCC 674  
 Qy 781 CTAAACGATGTCACTAGTTGTGGGCTTATAGGCTTGTGTATACGAAAGTCAACGCGTGA 840  
 Db 673 CTAAACGATGTCACTAGTTGTGGGCTTATAGGCTTGTGTATACGAAAGTCAACGCGTGA 614  
 Qy 841 AGTTGACCGCTGGGGAGTACGGTCCGCAAGATTAAACTCAAGAAATTGACGGGAGCC 900  
 Db 613 AGTTGACCGCTGGGGAGTACGGTCCGCAAGATTAAACTCAAGAAATTGACGGGAGCC 554  
 Qy 901 GCAACAGCGGTGATATATGTGATTAATTCGATGCAACGCGGAAACCTTACCTACCCCT 960  
 Db 553 GCAACAGCGGTGATATATGTGATTAATTCGATGCAACGCGGAAACCTTACCTACCCCT 494  
 Qy 961 GACATGTAGCGAATTTCTAGAGATAGATTAGTCT---TCGGGAAACGCTAACACAGGTG 1017  
 Db 493 GACATGTAGCGAATTTCTAGAGATAGATTAGTCT---TCGGGAAACGCTAACACAGGTG 434  
 Qy 1018 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
 Db 433 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374  
 Qy 1078 ACCCTTGTATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
 Db 373 ACCCTTGTATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319  
 Qy 1138 CCGAGAGAGGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGGGTAGGGCTTCAACG 1197  
 Db 318 CCGAGAGAGGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGGGTAGGGCTTCAACG 259  
 Qy 1198 TAATACATAGGCGGTCAGAGAGGTTGCCAACCCGAGGGGAGGCTAATCTCAAGAAAGC 1257  
 Db 258 TCATACATAGGCGGTCAGAGAGGTTGCCAACCCGAGGGGAGGCTAATCTCAAGAAAGC 199  
 Qy 1258 GCGTCTGATCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 1317  
 Db 198 GCGTCTGATCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 139  
 Qy 1318 CCGGATCAGCATGTGCGCGGTGAATAGTTCCTCCGCGTCTTGTACACACCGCCGTCACAC 1377  
 Db 138 CCGGATCAGCATGTGCGCGGTGAATAGTTCCTCCGCGTCTTGTACACACCGCCGTCACAC 79  
 Qy 1378 CATGGAGTGGGTTTTCACAGAGAGGTAAGTCTTAACCTGAAGAGGGCGCTTCCACGG 1437  
 Db 78 CATGGAGTGGGTTTTCACAGAGAGGTAAGTCTTAACCTGAAGAGGGCGCTTCCACGG 19  
 Qy 1438 TAGATTCATGACTGGGG 1455  
 Db 18 CAGGGTTCGTGACTGGGG 1  
 RESULT 67  
 ADG17998/C  
 ID ADG17998 standard; DNA; 1453 BP.  
 AC ADG17998;  
 XX  
 DT 26-FEB-2004 (first entry)  
 XX  
 DE Unknown bacterial rdna #4.  
 XX  
 KM 16S rdna; molecular biology; microbiology; unrecognised bacteria;  
 XX industrial wastewater; bioreactor; ds.  
 OS Unidentified.

XX  
 PN US2003203398-A1.  
 XX  
 PD 30-OCT-2003.  
 XX  
 PF 17-JUN-2003; 2003US-00464724.  
 XX  
 PR 16-DEC-1999; 99US-0171140P.  
 PR 13-DEC-2000; 2000US-00735567.  
 XX  
 PA (BRAM/) BRAMUCCI M G.  
 PA (CHEN/) CHEN M W.  
 PA (ALBE/) ALBERTSON-KANE H M.  
 PA (NAG/) NAGARAJAN V.  
 XX  
 PI Bramucci MG, Chen MW, Albertson-Kane HM, Nagarajan V;  
 DR WPI; 2003-906635/82.  
 XX  
 PT New isolated 16S rdna molecule, useful in the field of microbiology and  
 PT molecular biology, in particular for diagnosing organisms, specifically  
 PT previously unrecognized bacteria, from an industrial wastewater  
 PT bioreactor.  
 PS  
 XX  
 Claim 1; SEQ ID NO 4; 24bp; English.  
 CC  
 CC An isolated 16S rdna sequence (1) indicative of the presence of an  
 CC activated bacterial strain comprises: any of 7 fully defined sequences of  
 CC 1385-1539 bp (SEQ ID NO:1-7) or any of 15 fully defined sequences of 18-  
 CC 46 bp (SEQ ID NO:16-30), all given in the specification; hybridizing  
 CC with the sequence of (a) under conditions of 6x times; SSC (1 M NaCl), 50°  
 CC formamide, 1% SDS at 37degC, and a wash in 0.1x times; SSC at 60-65degC  
 CC ; at least 97% sequence identity to the sequence of (a); or complements  
 CC of the sequences of (a), (b) or (c). INDEPENDENT CLAIMS are also included  
 CC for the following: a method for identifying an activated sludge bacterial  
 CC strain, comprising extracting genomic DNA from a bacterial cell, and  
 CC probing the DNA with a probe derived from any of the sequences of (1) or  
 CC under suitable hybridization conditions, where the identification of a  
 CC hybridizable nucleic acid fragment confirms the presence of an activated  
 CC sludge bacterial strain, or amplifying the extracted DNA with at least  
 CC one oligonucleotide primer corresponding to a portion of any of the  
 CC sequences of (1) such that amplification products are generated, where  
 CC the presence of the amplification products confirms the presence of  
 CC activated sludge bacterial strain; a pyruvate degrading bacterial strain  
 CC having the 16S rdna sequence from SEQ ID NO: 1, 2, 5, 6 and 7; a acetate  
 CC degrading bacterial strain having the 16S rdna sequence from SEQ ID NO: 3  
 CC and 4; a benzoate degrading bacterial strain having the 16S rdna sequence  
 CC from SEQ ID NO: 2, 6 and 7; a denitrifying degrading bacterial strain  
 CC having the 16S rdna sequence from SEQ ID NO: 1 and 6; a method for the  
 CC degradation of pyruvate, acetate or benzoate, comprising contacting the  
 CC bacterial strain of (2)-(4) under suitable growth conditions with a  
 CC pyruvate, acetate or benzoate whereby the pyruvate, acetate or benzoate  
 CC is degraded, respectively; and a method for the reduction of nitrate,  
 CC comprising contacting the bacterial strain of (5) under suitable  
 CC conditions with a nitrate whereby the nitrate is reduced. Also disclosed  
 CC are polypeptides, vectors, host cells, primers, probes and a variety of  
 CC isolated strains used in the present invention. Preferred Method: The  
 CC bacterial cell in any of the methods cited above is contained within  
 CC wastewater sludge, and/or is contained within an environmental sample  
 CC selected from groundwater and soil. The methods and compositions of the  
 CC present invention are useful for molecular biology and microbiology,  
 CC particularly in identifying and isolating several previously unrecognized  
 CC bacteria from an industrial wastewater bioreactor. The isolation of  
 CC several strains of bacteria from a wastewater bioreactor was investigated  
 CC using several types of standard microbiological agar media. The 16S rdna  
 CC sequences of the bacterial strains were no more than 95% identical to the  
 CC 16S rdna sequences amplified and sequenced by standard PCR. 16S rdna  
 CC sequences that have less than 97% identity were most likely derived from  
 CC different bacterial species. Since the 16S rdna sequences derived from  
 CC EMB19A and NBY13 have less than 96% identity with the most similar 16S  
 CC sequences already present in the Genbank, the 16S rdna sequences from  
 CC EMB19A and NBY13 are significantly different from any other 16S rdna or  
 CC rdna sequence previously known to art.

XX Sequence 1453 BP; 279 A; 469 C; 344 G; 361 T; 0 U; 0 Other;  
SQ  
Query Match 75.5%; Score 1100; DB 10; Length 1453;  
Best Local Similarity 86.1%; Pred. No. 0;  
Matches 1256; Conservative 0; Mismatches 190; Indels 12; Gaps 3;  
1 ATTGAAGCTGGCGGCAATGCTTTACATATGCAATGCAACGGCAGACCGATGCTTGCAT 60  
1449 ATTGAAGCTGGCGGCAATGCTTTACATATGCAATGCAACGGCAGCGCGGCGCA----AC 1394  
QY CTGTGGCGAGTGGCGGCGGAGTGAATATCATTCGGAACGTAATCCGAAAGAGGGGGGTA 120  
DB CTGTGGCGAGTGGCGGCGGAGTGAATATCATTCGGAACGTAATCCGGAAGAGGGGGATTA 1334  
QY 121 ACGGATGAAGAGATGTGTAAATACCGATATCTAAGAGAGAAAGAGGGGATCGAAA 180  
DB 1333 GCCCGGGAAGCCGGAATTAATCCGATGTGATCTGAGAGATGAAGATGGGGGACCCGA 1274  
QY 181 GACCTTGGCTTTTGAAGCGCGCCGATGTCTGATTAGTGAATGGTGGGTTAAAGGCTTAC 240  
DB 1273 GGCTTCAAGGCTTTGAGAGCGCGCCGATGAGATTAGTGAATGGTGGGTTAAAGGCTTAC 1214  
QY 241 CAAGGCGACGATCACTGATTTGTCTGAGAGAGACGACCGACCACTGGGATCTGAGACAG 300  
DB 1213 CAAGCTGCGATCTGTACTGTCTGTGAGAGATGATCAAGCCACTGGGACTGAGACACG 1154  
QY 301 GCCGAGATCTCTAAGGGGAGGAGCAGTGGGGAAATTTTGAACATGGGCGCAAGCCTGATC 360  
DB 1153 GCCGAGATCTCTAAGGGGAGGAGCAGTGGGGAAATTTTGAACATGGGCGCAAGCCTGATC 1094  
QY 361 CAGCAATGCCCGCTGATGAGAAAGGCGCTTGGTTGTTAAAGCTTTTCACTGAGAGAGA 420  
DB 1093 CAGCAATGCCCGCTGATGAGAAAGGCGCTTGGTTGTTAAAGCTTTTGTACGGAAGA 1034  
QY 421 AAAGGTTACGTTAAATTAATCTGTACCCATGACCGTATGACAGAAAGACCGGCTTAC 480  
DB 1033 AAAGGCTCTCTCTATATACAGGGGGCATATGACGGTACGTTAAGATTAAGACCGGCTTAC 974  
QY 481 TAAGTCCAGAGAGCGCGGTTAATAGTAGGGTGAAGGCTTAATCGAATTAATCTGGGCGT 540  
DB 973 TAAGTCCAGAGAGCGCGGTTAATAGTAGGGTGAAGGCTTAATCGAATTAATCTGGGCGT 914  
QY 541 AAAGGTTACGAGCGCGCTTGTAAAGTGAATGTTAAATCCCGGCGCTTAACTTGGGAT 600  
DB 913 AAAGGTTACGAGCGCGCTTGTAAAGTGAATGTTAAATCCCGGCGCTTAACTTGGGAT 854  
QY 601 TGGCTTTGAAACTCAAAAGTGAAGTGGCAGAGGAGTGGAAATTCATGTGTGACAG 660  
DB 853 TGGCTTTGTAAGTGAAGGCTGAGTGGCAGAGGAGGAGTGGAAATTCGCTGTGACAG 794  
QY 661 TGAATGGTGAAGATATGGAAGACATCGATGGGGAAGGCGCTCCGCGGTTAAACT 720  
DB 793 TGAATGGTGAAGATATGGAAGACATCGATGGGGAAGGCGCTCCGCGGCTTAAACT 734  
QY 721 GACGCTCATGCAAGAAAGCTGGGAGCAACAGGATTAAGATCCCTGGTATCCACGCC 780  
DB 733 GACGCTCATGCAAGAAAGCTGGGAGCAACAGGATTAAGATCCCTGGTATCCACGCC 674  
QY 781 CTAAACGATGTCACTAGTTGTTGGGCTTATTAAGCTTGGTGAACGAAGCTTAACCGCTGA 840  
DB 673 CTAAACGATGTCACTAGTTGTTGGGAAATTCATTTTCTCAGTAAAGAGCTTAACCGCTGA 614  
QY 841 AGTTAAGCGGCTGGGAGAGTGGCAAGATTAAACCTCAAGGAATTGACGGGGAAGCC 900  
DB 613 AGTTAAGCGGCTGGGAGAGTGGCAAGATTAAACCTCAAGGAATTGACGGGGAAGCC 554  
QY 901 GCACAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAAACTTACCTTACCTT 960  
DB 553 GCACAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAAACTTACCTTACCTT 494  
QY 961 GACATGTACGGAATTTCTAAGATAGATTAGTGCT--TCGGGAACGCTTAACACAGGTG 1017

DB 493 GACATGTACGGAATTCCTGACAGAGACGACAGAGTGTCTGAAGAAGACCGCTTAACACAGGTG 434  
QY 1018 CTGCATGCTGTGTGTCACTCTGTGTGTGAGATGTTGGGTTAAGTCCCGAAGAGCGCA 1077  
DB 433 CTGCATGCTGTGTGTCACTCTGTGTGTGAGATGTTGGGTTAAGTCCCGAAGAGCGCA 374  
QY 1078 ACCCTTGTCAATTAATTTGCCATCAATTTGGTTGGGCACTTTAATGAGACTGCGGTTGACAA 1137  
DB 373 ACCCTTGTCAATTAATTTGGCTTACGAAA-----GGGCACTTAATGAGACTGCGGTTGACAA 319  
QY 1138 CCGAGGAAGGTGGGATGACGTCAGATCTCATGAGCCCTTAATGAGGCTTACACAG 1197  
DB 318 CCGAGGAAGGTGGGATGACGTCAGATCTCATGAGCCCTTAATGAGGCTTACACAG 259  
QY 1198 TAATACATGCGCGCTTACAGAGGTTGGCCAAACCCCGAGGGGAGCTAATCTGAGAAAGC 1257  
DB 258 TCATACATGCTGTGTGACAAAGGGTTGCCAAACCCCGAGGGGAGCTAATCTGAGAAAGC 199  
QY 1258 GCGTCTGATCCGGATCGGATCGACACTGCACTCGATCGGAATCGGATCGCTAGTAAT 1317  
DB 198 GCGTCTGATCCGGATCGGATCGACACTGCACTCGATCGGAATCGGATCGCTAGTAAT 139  
QY 1318 CCGGATCAGCATGTCTCGGTTGAATAGTTCCCGGCTTTGTACACACCGCCGTCACAC 1377  
DB 138 CCGGATCAGCATGTCTCGGTTGAATAGTTCCCGGCTTTGTACACACCGCCGTCACAC 79  
QY 1378 CATGGAGTGGGTTTACCGAAGCAGGTAGTCTAACCGTAAGAGGGCGCTTGGCCACGG 1437  
DB 78 CATGGAGTGGGTTTCCGCAAGAGTAGTTAGCTTAACCGAAGAGGGCGATTACACAGG 19  
QY 1438 TGAGATTCAATGCTGGGG 1455  
DB 18 CAGGTTCTGTACTGGGG 1

RESULT 68  
ADL27933/C  
ID ADL27933 standard; DNA, 1453 BP.  
XX  
XX ADL27933;  
AC  
XX 20-MAY-2004 (first entry)  
DT  
XX  
DE RA2 16S rDNA.  
XX  
XX Activated sludge bacterial strain; pyruvate; degrading bacterial strain;  
KW acetate; benzoate; nitrate; wastewater sludge; groundwater; soil;  
KW nitrate reduction; RA2; 16S rDNA; de.  
OS  
XX  
XX Unidentified.  
PN US2003207320-A1.  
XX  
XX 06-NOV-2003.  
XX  
XX 17-JUN-2003; 2003US-00464356.  
PF  
XX 16-DEC-1999; 99US-0171140P.  
PR 13-DEC-2000; 2000US-00735567.  
XX  
XX (BRAM/) BRAMUCCI M. G.  
PA (CHEN/) CHEN M. W.  
PA (ALBE/) ALBERTSON-KANE H. M.  
PA (NAGA/) NAGARAJAN V.  
PI  
XX Brumucci MG, Chen MW, Albertson-Kane HM, Nagarajan V;  
XX WPI; 2003-901055/82.  
DR  
XX  
XX New isolated 16S rDNA sequence indicative of the presence of an activated  
PT sludge bacterial strain, useful for rapidly identifying and monitoring  
PT bacterial strains in samples taken from a wastewater bioreactor.

PS Claim 1; SEQ ID NO 4; 25pp; English.  
CC The invention relates to an isolated 16S rDNA sequence. The isolated 16S  
CC rDNA sequence is indicative of the presence of an activated sludge  
CC bacterial strain. The invention also relates to methods for identifying  
CC an activated sludge bacterial strain, a pyruvate degrading bacterial  
CC strain, an acetate degrading bacterial strain, a benzoate degrading  
CC bacterial strain, methods for degrading pyruvate, acetate or benzoate and  
CC a method for reducing nitrate. Identifying an activated sludge bacterial  
CC strain comprises extracting genomic DNA from a bacterial cell and probing  
CC the extracted genomic DNA with a probe derived from any of the above-  
CC mentioned sequences under suitable hybridization conditions, where the  
CC identification of a hybridizable nucleic acid fragment confirms the  
CC presence of an activated sludge bacterial strain. Alternatively, the  
CC method comprises extracting genomic DNA from a bacterial cell and  
CC amplifying the extracted genomic DNA with at least one oligonucleotide  
CC primer corresponding to a portion of the above-mentioned sequences, such  
CC that amplification products are generated, where the presence of  
CC amplification products confirms the presence of an activated sludge  
CC bacterial strain. The bacterial cell is contained within wastewater  
CC sludge or within an environmental sample selected from groundwater and  
CC soil. Degrading pyruvate, acetate or benzoate comprises contacting the  
CC bacterial strain under suitable growth conditions with an amount of  
CC pyruvate, acetate or benzoate. Reducing nitrate comprises contacting the  
CC bacterial strain under suitable growth conditions with an amount of  
CC nitrate. The 16S rDNA sequences and methods are useful in rapidly  
CC identifying and monitoring the corresponding bacterial strains in samples  
CC taken from a wastewater bioreactor or from any environment which might  
CC contain the specified bacteria. This sequence represents an isolated 16S  
CC rDNA of the invention.  
XX  
SQ Sequence 1453 BP; 279 A; 469 C; 344 G; 361 T; 0 U; 0 Other;  
Query Match 75.5%; Score 1100; DB 11; Length 1453;  
Best Local Similarity 86.1%; Pred. No. 0;  
Matches 1256; Conservative 0; Mismatches 190; Indels 12; Gaps 3;  
QY 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAAGTCGAACGCGACGACGCGATGCTTGCAT 60  
DB 1449 ATTGAACGCTGGCGGCGATGCTTTACATGCAAGTCGAACGCGACGCGCGGCGCA---AC 1394  
QY CTGGTGGCGAGTGGCGGACGGGTAGATGATCGGAAGCTATCCAGAAGAGGGGGATA 120  
DB 1393 CTGGCGGCGAGCGGCGCAACGGGTAGATGATCGGAAGCTATCCAGAAGAGGGGGATA 1334  
QY 121 ACGCATCGAAGATGTGTCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATCGAA 180  
DB 1333 GCCCGGCGAAGCGCGATTAATACCGCATGATCTGAGGATGAAAGTGGGGACCGCAA 1274  
QY 181 GACCTTGGCTTTTGAAGCGCGGATGTCTGATTAAGTATGTTGGGGTAAAGGCTTAC 240  
DB 1273 GGCGCTCAACGGCTTGGAGCGCGCGATGAGATTAAGTATGTTGGGGTAAAGGCTTAC 1214  
QY 241 CAAGGCGAGATCAGTATGTTGTCTGAGAGACGACGACGACCACTGGGACTGAGACG 300  
DB 1213 CAAGCTTGGCATCTGTAGCTGTGTGAGAGATGATACGCACTGGGACTGAGACG 1154  
QY 301 GCCCAGACTCTCTACGGGAGCGACGATGGGAAATTTTGGACAATGGGCGGCAAGCTGATC 360  
DB 1153 GCCCAGACTCTCTACGGGAGCGACGATGGGAAATTTTGGACAATGGGCGGCAAGCTGATC 1094  
QY 361 CAGCAATGCCGCGGTGAGTGAAGAGCCCTTGGGTTGAAGCTCTTTCACTGAGAGAA 420  
DB 1093 CAGCCATTTCCGCGTGCAGAGCAAGGCCCTTGGGTTGAAGCTCTTTTATCGGAACGA 1034  
QY 421 AAAGGTTAAGGTAAATATATGTGACCGATGACGCTATCGACAGAAAGAGCAGCGGCTAAC 480  
DB 1033 AAAGGCTCTCTCTTAATCAAGGGGCAATATGACGCTATCGTAAGATTAAGCAGCGCTAAC 974  
QY 481 TAGGTGCACAGCCGCGGTAAATATGTAAGGTGCAACGTTAATCGGAATTAATCTGGGCGT 540  
DB 973 TAGGTGCACAGCCGCGGTAAATATGTAAGGTGCAACGTTAATCGGAATTAATCTGGGCGT 914

QY 541 AAAGGTGCGACGCGCGCTTGTAAATGATGTGAATCCCGGCGCTTAACCTGGGAAT 600  
DB 913 AAAGGTGCGACGCGCGCTTGTAAATGATGTGAATCCCGGCGCTTAACCTGGGAAT 854  
QY 601 TGGCTTTGAACCTCAAAAGCTTAAGTGGCGACAGAGAGTGGAAATTCATGTGACAG 660  
DB 853 TGGCTTTGATGACGAAAGCTGAGTGGCGGACAGAGAGAGTGGAAATTCGCGCTGACAG 794  
QY 661 TGAATGTCTAAGATATGTGAAGAAATCATGATGGCGGAAGGCGCTCCGCGGTTAACT 720  
DB 793 TGAATGTCTAAGATATGTGAAGAAATCATGATGGCGGAAGGCGCAATCCCTGGGCTGCACT 734  
QY 721 GACGCTCATGACGAAAGGCTGGGAGCAAAAGGATTAGATACCTGGTATGTCAGCGC 780  
DB 733 GACGCTCATGACGAAAGGCTGGGAGCAAAAGGATTAGATACCTGGTATGTCAGCGC 674  
QY 781 CTAAACGATGTCAACTAGTGTGGGCTTATTAAGCTTGTGTAACGACCTTAACGCTGA 840  
DB 673 CTAAACGATGTCAACTAGTGTGGGATTTCAATTTCTCATGTAACGAACTTAACGCTGA 614  
QY 841 AGTTGACCGCTGGGAGATACGCTGCGCAAGATTTAAATCTAAAGAAATGACGGGAGCC 900  
DB 613 AGTTGACCGCTGGGAGATACGCTGCGCAAGATTTAAATCTAAAGAAATGACGGGAGCC 554  
QY 901 GCACAAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAGAACTTAACCTTACCCT 960  
DB 553 GCACAAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAGAACTTAACCTTACCCT 494  
QY 961 GACATGTACGCAATTTCTAGAGATATGATTAAGTCT--TCGGGAAAGCTTAACACAGGTG 1017  
DB 493 GACATGTACGCAATCTGACAGAGACGACAGAGATGCTCGAAGAGAGCGTAAACACAGGTG 434  
QY 1018 CTGCATGCTGTGCTGTAGTCTGCTGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCA 1077  
DB 433 CTGCATGCTGTGCTGTAGTCTGCTGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCA 374  
QY 1078 ACCCTGTCAATTAATGGCATCATTTGTTGGGCACTTTAATGAGACTCCGGTGAAGAA 1137  
DB 373 ACCCTGTCAATTAATGGCTTCAACGAAA-----GGGCACTTAATGAGACTCCGGTGAAGAA 319  
QY 1138 CCGGAGAAAGTGGGAGATGACGTAAGTCTTATGAGCTTATGGGTAGGGCTTACACG 1197  
DB 318 CCGGAGAAAGTGGGAGATGACGTAAGTCTTATGAGCTTATGGGTAGGGCTTACACG 259  
QY 1198 TAATTAATAGGCGGCTACAGAGGCTTCCAAACCGCGAGGGGAGCTAATCTCAGAAAG 1257  
DB 258 TCATTAATAGGCTGCTGTAACAAAGGCTTCCAAACCGCGAGGGGAGCTAATCTCAGAAAG 199  
QY 1258 GCGTGTAGTCCGATCGGAGTCTGCACTGCACTCCGTGAAGTCCGAATCGCTAGTAAT 1317  
DB 198 GCGTGTAGTCCGATCGGAGTCTGCACTGCACTCCGTGAAGTCCGAATCGCTAGTAAT 139  
QY 1318 CCGGATCAGCATGTGCGGCTGAATACGTTCCCGGCTTTGTACACACCGCCGCTACAC 1377  
DB 138 CCGGATCAGCATGTGCAAGTGAATCGTTCCCGGCTTTGTACACACCGCCGCTACAC 79  
QY 1378 CATGGAGTGGGTTTACCAAGAAAGCATAGTCTAACCGTAAAGAGGGGCTTGGCCACG 1437  
DB 78 CATGGAGTGGGTTTCCCAAGAAAGTATGATCTTAACCGCAAGAGAGGGGATTAACACG 19  
QY 1438 TGAGATTCAATGAGGCG 1455  
DB 18 CAGGGTTCTGACTGGGG 1  
RESULT 69  
ADP47789/c  
ID ADP47789 standard; DNM: 1453 BP.  
XX  
XX ADP47789;  
AC  
AC  
DT 12-FEB-2004 (first entry)  
XX

DE Unknown bacterial strain RA2 16S rDNA.  
XX  
XX 16S rDNA; activated sludge; bacterial strain; diagnosing organism;  
KM unrecognised bacteria; industrial wastewater bioreactor; db.  
XX  
XX Unidentified.  
OS  
XX US6608190-B1.  
XX  
XX 19-AUG-2003.  
PD  
XX 13-DEC-2000; 2000US-00735567.  
PF  
XX 16-DEC-1999; 99US-0171140P.  
PR  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
XX Bramucci MG, Chen MW, Albertson-Kane HM, Nagarajan V;  
PI WPI; 2004-008386/01.  
XX  
XX  
XX New isolated 16S rDNA molecule, useful in the field of microbiology and  
PT molecular biology, in particular for diagnosing organisms, specifically  
PT unrecognised bacteria, from an industrial wastewater bioreactor.  
XX  
XX Disclosure; SEQ ID NO 4; 43pp; English.  
PS  
XX The invention relates to an isolated 16S rDNA molecule indicative of the  
CC presence of an activated sludge bacterial strain. The compositions of the  
CC present invention are useful in the field of microbiology and molecular  
CC biology, in particular for diagnosing organisms, specifically  
CC unrecognised bacteria, from an industrial wastewater bioreactor. The  
CC present sequence is used in the exemplification of the invention.  
XX  
XX Sequence 1453 BP; 279 A; 469 C; 344 G; 361 T; 0 U; 0 Other;  
SQ  
Query Match 75.5%; Score 1100; DB 12; Length 1453;  
Best Local Similarity 86.1%; Pred. No. 0;  
Matches 1256; Conservative 0; Mismatches 190; Indels 12; Gaps 3;  
QY 1 ATTGAAGCTTGGCGGCGATCTTTACATGCAAGTCCGACGGGCGACGATGCTTGAT 60  
DB 1449 ATTGAAGCTTGGCGGCGATCTTTACATGCAAGTCCGACGGGCGCA---AC 1394  
QY 61 CTGGTGGCGAGTGGCGGCGAGTATGATCATCGGACGATCTATCCAGAAAGGGGGGTA 120  
DB 1393 CTGGCGGCGAGCGGCGAGCGGTAAGTACACATCGAACCTGCGGCGGCGATA 1334  
QY 121 ACGCATGMAAGATGTGCTAATACCGCATATATCTTAAGAGAAAGCGGGGATCGAA 180  
DB 1333 GCCCGGGGAAAGCGGATTAATACCGCATGTGATGAGATGAAGTGGGGGACCGCA 1274  
QY 181 GACCTTGGCTTTTGGAGCGGCGGATGTCTGATTAGCTAAGTTGGGGGTAAAGGCTTAC 240  
DB 1273 GGCTTCAAGCTTTGGAGCGGCGGATGCAATTAAGTATGGGGGTAAAGGCTTAC 1214  
QY 241 CAAGGCGAGATCGTATGTTGCTGAGAGACGACGACACTGGGATCGAGACAG 300  
DB 1213 CAAGCTGCGATCTGTACTGCTGTGAGAGATGATCCACACTGGGATCGAGACAG 1154  
QY 301 GCCGAGACTCTTACGAGGAGGAGACAGTGGGAAATTTGGACATATGGGCGCAAGCTGATC 360  
DB 1153 GCCGAGACTCTTACGAGGAGGAGACAGTGGGAAATTTGGACATATGGGCGCAAGCTGATC 1094  
QY 361 CAGCAATGCCGCTGATGAGAGAGGCTTGGGTTGTAAGCTTTTCACTCGAGAGA 420  
DB 1093 CAGCAATTCGCGCTCAGAGACGAAAGCCTTGGGTTGTAAGCTTTTCACTCGAGAGA 1034  
QY 421 AAAGGTTACGTAATTAATCGGACCCATGACGATGACAGAGAGACACCGGCTTAC 480  
DB 1033 AAAGGCTCTCTTAATACAGGGGCAATGACGGATACGTAAGTAAGACACCGGCTTAC 974  
QY 481 TACGTGCAAGACCGCGGTAATACGTAGGCTGCAAGCGTTAATGGAATTACTGGGCGT 540

DB 973 TACGTGCAAGACCGCGGTAATACGTAGGCTGCAAGCGTTAATGGAATTACTGGGCGT 914  
QY 541 AAAGGTTGCGAGCGGCGCTTGAATGATGATGTAATCCCGGCTTAATCTGGGAAT 600  
DB 913 AAAGGTTGCGAGCGGCTTGAATGATGATGTAATCCCGGCTTAATCTGGGAAT 854  
QY 601 TGGCTTTGAAATCTCAAAACCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGAG 660  
DB 853 TGCCTTTGACCTGCAAGGCTGAGTGCAGAGAGGAGATGAATTCGCGGTGTAGAG 794  
QY 661 TGAATGCTGAGATATGAGAAACATGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
DB 793 TGAATGCTGAGATATGCGAGAGGAGACCGATGGGAGGAGGAGGAGGAGGAGGAGGAG 724  
QY 721 GACGCTCATGACGAAAGCGTGGGAGGAGCAACAGATTGATTCCTGTTGATGACGCG 780  
DB 733 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTGATTCCTGTTGATGACGCG 674  
QY 781 CTAAAGATGTCACTAGTTGTTGGGCTTATTAAGGCTTGTAAAGGAGTAAAGGAGTGA 840  
DB 673 CTAAAGATGTCACTAGTTGTTGGGATTCATTTCTCAATGAGAGGAGTAAAGGAGTGA 614  
QY 841 AGTTGACCGCTGGGAGGAGTACGGTGCAGAGATTAAACTCAAGAGATTGACGGGAGCC 900  
DB 613 AGTTGACCGCTGGGAGGAGTACGGGCGGAGGTTAAATCAAGAGATTGACGGGAGCC 554  
QY 901 GCACAAGCGGTGATTAATGATGATTAATTCATGATGCAACGGGAAACCTTACCTACCTT 960  
DB 553 GCACAAGCGGTGATTAATGATGATTAATTCATGATGCAACGGGAAACCTTACCTACCTT 494  
QY 961 GACATGAGCGAATTTTCTAGATGATTAATGATGATGATGATGATGATGATGATGATGAT 1017  
DB 493 GACATGAGCGAATTCCTGACAGACGCAAGAGGCTCGAAGAGGCGGATGACAGGATG 434  
QY 1018 CTGCATGAGCGTGTGTCAGCTGATGTCGTGATGATGATGATGATGATGATGATGATGATGAT 1077  
DB 433 CTGCATGAGCGTGTGTCAGCTGATGTCGTGATGATGATGATGATGATGATGATGATGATGAT 374  
QY 1078 ACCCTTGTCTAATTAATGCTCATATTTGGTGGGCACTTAAATGAGTCCCGGTGACAA 1137  
DB 373 ACCCTTGTCTAATTAATGCTCATATTTGGTGGGCACTTAAATGAGTCCCGGTGACAA 319  
QY 1138 CCGAGAGAGGTTGGGATGACGTCAGTCTCATGCGGCTTAATGAGGAGGCTTCAACAG 1197  
DB 318 CCGAGAGAGGTTGGGATGACGTCAGTCTCATGCGGCTTAATGAGGAGGCTTCAACAG 259  
QY 1198 TAAATCAATGGCGGTACAGAGGTTGCCAACCAGGAGGAGGAGGAGGAGGAGGAGGAGG 1257  
DB 258 TCATACATAGCTGTGTAACAAAGGTTGCCAACCAGGAGGAGGAGGAGGAGGAGGAGGAGG 199  
QY 1258 GCGTGTAGTCCGATCGAGTCTGCACTCGATCGATCGATCGATCGATCGATCGATCGAT 1317  
DB 198 CAGTGTAGTCCGATCGAGTCTGCACTCGATCGATCGATCGATCGATCGATCGATCGAT 139  
QY 1318 CCGGATCAAGTGTCCGTTGAATACGTTCCGAGTCTTGTACACACCGGCTGACAC 1377  
DB 138 CCGGATCAAGTGTCCGTTGAATACGTTCCGAGTCTTGTACACACCGGCTGACAC 79  
QY 1378 CATGGAGTGGGTTTCAACGAAAGCAGTGTCTTAACGTAAGGAGGAGGAGGAGGAGGAGG 1437  
DB 78 CATGGAGTGGGTTTCAACGAAAGTGTAGCTTAACGTAAGGAGGAGGAGGAGGAGGAGG 19  
QY 1438 TGAATTCATGACTGGG 1455  
DB 18 CAGGTTCTGTACTGGG 1  
RESULT 70  
AAA60564  
ID AAA60564 standard; DNA; 1387 BP.  
XX  
AC AAA60564;







Tel: 33 1 34 65 25 16  
Fax: 33 1 34 65 25 21  
Email: sorokine@jouy.inra.fr  
best homologue in strain IL1403 is ywga (78%)  
Class: shotgun  
High quality sequence start: 30  
High quality sequence stop: 6471.  
Location/Qualifiers

## FEATURES

source

1. 6499  
/organism="Lactococcus lactis subsp. cremoris"  
/mol\_type="genomic DNA"  
/strain="MG1363"  
/sub\_species="cremoris"  
/DB\_xref="taxon:1359"  
/clone\_lib="MG1363 Random Sequence Tag Library"  
/note="Vector: pSGM02; Site 1: SmaI; Library of  
chromosomal fragments of L. lactis strain MG1363 was  
prepared by partial AluI digestion or by sonication."

## ORIGIN

Query Match 55.6%; Score 810.6; DB 9; Length 6499;  
Best Local Similarity 75.3%; Pred. No. 1,7e-238;  
Matches 1104; Conservative 0; Mismatches 349; Indels 14; Gaps 7;  
4 GAACGCTGGGCGATGCTTTACATGCAAGTGAAG--GCAGCAGGATGCTTGATC 61  
555 GAACGCTGGGCGGCTGCTAATCATGCAAGTGAAGATTTGGCTTGACCA 614  
62 TGTGGCGAGTGGCGGAGCGGTGATGATCAT--CGAAGCTATCCAGAGAGGGGTA 120  
615 ATTGGAAGGACAGCAAGCGGTGATGACGCGTGGGAATCTGCTTTGACGGGGGACA 674  
121 AGCGATGAAAGATGCTAATACCGCAT-----TACTTAAGAGAAAGCAAGGG 173  
675 ACATTGGAAGCAAGATGCTAATACCGCATATTAACATTAAAGTTTAAGTTGA 734  
174 ATGGAAGACCTTGGCTTTGAGCGGCGGATGCTGATTAAGTCTGGTGGGTTAA 233  
735 AGATGCAATTGCATCACTCAAAAGATGATCCGGGTGTATTAAGTGTGTAAGGTA 794  
234 GGCCTACCAAGGCGAGCATCATGATGATGCTGAGAGAGACAGCAGCAGCACTGGGACTG 293  
795 GGGTCCACCAAGGCGAGTATCATACCGACCTGAGAGGGTATCGGCAATGGGACTG 854  
294 AGACACGGCCAGACCTCTACGAGAGGACAGTGGGAAATTTGACAAATGGCGGCAAG 353  
855 AGACACGGCCCAAACTCTACGAGAGGACAGTAAAGGAATCTTCGCAATGACGAAAG 914  
354 CCGATTCAGCAATGCCCGCTGAGTGAAGAGCCTTCGGGTTGTAAAGCTCTTCAATC 413  
915 TCTGACCGAGCAACGCCCGCTGAGTGAAGAGGTTTCGGATCGTAAACCTCTGTGGTA 974  
414 GAGAGAAAGGTTACGGTAAATATGATGCCA--TGACGATATCGACAGAAAGAC 472  
975 GAGAGAAAGCTGTGAGATGAGAAAGCTCATCAAGTGAAGCGTAACTACCAAGAAAGGA 1034  
473 CGGCTAATCTAGTCCAGCAGCGCGGTAAATACGTAGGGTCAAGCGTTAATGGAATTA 532  
1035 CGGCTAATCTAGTCCAGCAGCGCGGTAAATACGTAGGGTCCCAAGCCTTGTCCGGAATTA 1094  
533 CTGGGCTTAAAGGTGGCGAGCGGCTTGAAGTCAAGTGAATCCCGGGGCTTAAC 592  
1095 TTGGGGGTAAAGGAGGAGCGGTGTTTATTAAGTCTGGGTAAAGGCGTGGCTCAAC 1154  
593 CTGGGATTTGCTTGAATCAAAAGCTAGAGTGTGGCAGAGGAGGTGAATTCATG 652  
1155 CATTTGA--TGCATTTGAGAACTGTAGCTTAGTGTGAGAGAGAGAGTGAATTCATG 1213  
653 TGTAGCAGTGAATGCTGATGATATGAGAAAGCAATCATGATGCGAGAGCAGCTCTCGG 712  
1214 TGTAGCGGTGAATGCTGATATATGAGAAAGCAACCGGTGCGAAAGCGGCTCTCGG 1273  
713 TTAACTGACGCTCATGCAAGAAAGGTGGGAGCAAAAGATTAAGATACCTGGTAG 772

DB 1274 CTGTACTGACACTGAGGCTCGAAACCGTGGGAGCAACAGATTAAGTACCTGGTAG 1333  
773 TCCAGCGCTTAACAGATGTCAACTAGTGTGGGCTTATTTG--GCTTGATACGAAGCT 831  
1334 TCCAGCGCTTAACAGATGTCAGTGTAGTGTAGGAGCTAATTAAGTTCTGTATCGAGCT 1393  
832 AACGCGTGAAGTTGACCGGCTGGGGAGTACGCGTCCGAAGATTAACCAAGAAATTA 891  
1394 AACGCAATAGACTCTCGCTGGGAGTACGACCCGAAAGTTGAATCAAGAAATTA 1453  
892 CGGGAGCCCGCACAAGCGGTGATTAATGATTAATTCATGCAACCGGAAACCTTA 951  
1454 CGGGGGCCCGCACAAGCGGTGAGATGATGTTAATGAGCAACCGGAAACCTTA 1513  
952 CCTACCTTTGACAT--GTACGAAATTTTCTAGAGATGATTAAGTCTTGGGAAAGCTAAC 1010  
1514 CAGGCTTTGACATCTCTGTCTATTCCTAGAGATGAGAAAGTTCTTGGGCAACGGGAT 1573  
1011 ACAGGTCGTCAGTGGCTGTCAGCTGTCGAGATGTCGAGATGTTGGTTAAGTCCGCAAC 1070  
1574 ACAGGTCGTCAGTGGCTGTCAGCTGTCGAGATGTCGAGATGTTGGTTAAGTCCGCAAC 1633  
1634 GAGCGCAACCTTATTTGATTAAGTTCATCATTAAGTGGGCACTTAACGAGACTGCGG 1693  
1071 GAGCGCAACCTTGTCTAATTAATTCGATCATTTGGTGGGCACTTAATGAGACTGCGG 1130  
1131 TGACAAACCGAGAGAGGTTGGGATGATCAAGTCTCATATGCGCTTATGGGTAGGCT 1190  
1694 TGTATTAACCGAGAGAGGTTGGGATGATCAAGTCTCATATCATATCCCTTAATGACCTGGGCT 1753  
1191 TGACAGTAAATCAATGGCGGCTGACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTC 1250  
1754 ACACAGTCTCAAGTGAAGTGTACACAGATCGGACAGATGATTTACTAATCTC 1813  
1251 AGAAAGCGGCTGATCCGATTCGAGTCTGCAACTCGACTCCGTGAAGTGGAAATCGC 1310  
1814 TTAACCACTTCTCAGTTCGATTTGAGGCTCAACTGCTCAATGAAGTGGAAATCGC 1873  
1311 TGAATTCGCGGATGACATGTCGCGGTGAATAGTCCCGGCTTTGTACACCGGCC 1370  
1874 TAGTAATCGCGATCGACAGCGCGGTGAATAGTCCCGGCTTTGTACACCGGCC 1933  
1371 GTCAACCAATGGAAGGTTTCAACAGAGAGGATGCTAACCGTAAAGAGGGGCTT 1430  
1934 GTCAACCAAGGAGTGGAGTACCCGAAGTGGTTGCTTAACCGCAAGAGGGGCTT 1993  
1431 GCCACGATGATTCATGACTGGGGTG 1457  
1994 CTTAAGTAAAGCAATGACTGGGGTG 2020  
DB  
RESULT 2  
CX109203 1954 bp mRNA linear EST 03-JUN-2005  
LOCUS RECM0593 A normalized whole-life-cycle cDNA library of rice Oryza  
DEFINITION sativa (indica cultivar-group) cDNA clone E1051022, E105108,  
E1008B17, B1076L13, B1154A06, B113B10, E111 5', mRNA sequence.  
ACCESSION CX109203  
VERSION CX109203.1 GI:66922355  
KEYWORDS EST.  
SOURCE Oryza sativa (indica cultivar-group)  
ORGANISM Oryza sativa (indica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriaristidae; Oryzaceae; Oryza.  
1 (bases 1 to 1954)  
Zhang, J., Feng, Q., Jin, C., Qiu, D., Zhang, L., Xie, K., Yuan, D.,  
Han, B., Zhang, Q., and Wang, S.  
Features of the expressed sequences revealed by a large-scale  
analysis of ESTs from a normalized cDNA library of the elite indica  
rice cultivar Minghui 63  
JOURNAL Plant J. 42 (5), 772-780 (2005)



PUBMED 1591889  
Contact: Wang S  
National Key Laboratory of Crop Genetic Improvement  
Hauzhong Agricultural University  
Wuhan 430070, China  
Tel: 86-27-87282044  
Fax: 86-27-87287092  
Email: shiyingwang@hotmail.com  
Seq primer: T7.

FEATURES  
Source Location/Qualifiers  
1..1954  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="mRNA"  
/strain="indica"  
/cultivar="Minghui 63"  
/db\_xref="taxon:39946"  
/clone="E1051022, E1053008, E1008B17, B1076L19, B1154A06, B113B10, B111"  
/issue\_type="whole plant"  
/dev\_stage="whole-life-cycle"  
/lab\_host="E. coli DH10B"  
/clone\_lib="A normalized whole-life-cycle cDNA library of rice"  
/note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; The library is constructed based on the strategy of saturation hybridization with genomic DNA using rice cultivar Minghui 63. This library consists of cDNA from 15 directionally cloned cDNA libraries constructed with different tissues from 9 developmental stages."

## ORIGIN

Query Match 49.6%; Score 722; DB 8; Length 1954;  
Best Local Similarity 73.2%; Pred. No. 3,1e-211;  
Matches 1070; Conservative 0; Mismatches 335; Indels 57; Gaps 9;

```
3 TGAACGCTGGCGGATGCTTTACATGCAAGTGAACGCGACGAGCTTGCATCT 62
130 TGAACGCTGGCGGATGCTTTACATGCAAGTGAACGCGGAGCTGT----- 177
63 GGTGGCGAGTGGCGGACCGGCTGAATGCAATCGGAACGTATCGAAGAGGGGGTAC 122
178 -GTTTCCAGTGGCGAAGCGGTGAATGCGCTAAGAACTGCTTGGAGAGGAAACAC 236
123 GCATCGAAGATGCTTAATACCGCATATCTCTAAGAGGAAAGAGGGGATCGAAGA 182
237 AACTGGAACCGGTGCTTAATACCGCATAGCT-----GAGGACAAAGAG 283
183 CTTTGGCGTTTGAAGCGCGGATGCTGATTAAGTGAATGCTGAGGCTTAACCA 242
284 AATTCGCGCCAAAGAGGGGCTCGGCTGATTAAGTGAATGCTGAGGCTTAACCA 343
243 AGGCGAGATCAATGATGCTGCTGAAGAGCAACGACACAGTGGGCTGAACAAGGC 302
344 AGGCGAGATCAATGATGCTGCTGAAGAGCAACGACACAGTGGGCTGAACAAGGC 403
303 CCAGACTCTTACCGGAGGCGACAGTGGGGAATTTTGAACAATGGCGCAACCTGATCA 362
404 CCAAGCTCTTACCGGAGGCGACAGTGGGGAATTTTCCCAATGGCGCAACCTGACGA 463
363 GCAATGCGCGGTGAATGAAGAGGCTTGGGGTGTAAAGCTCTTTCAGTGAAGAA 422
464 GCAATGCGCGGTGAATGAAGAGGCTTGGGGTGTAAAGCTCTTTCAGTGAAGAA 523
423 AGGTTACGCTAATATGCTGACCATGACGCTAATCGACAGAAAGACGCGCTAATCA 482
524 -----CAATGACGCTAATCGACAGAAATGACATGCGCTAATCTC 560
483 CGTGGCAGAGCGCGGTAAATCAATGAGGTGAACGCTTAATCGAATTAATGAGCGCTAA 542
561 TGTGCAAGACCGCGCTAAGCAAGATGCAACGCTTAATCGAATTAATGAGCGCTAA 620
543 AGGTTGCGCAGCGCGCTTGTAAAGTGAATGCAATCGCGGCTTAATCGGGAATTG 602
```

```
DB 621 AGGCTCTGAGGAGGCTTTTCAAGTCCGCGCTCAATCCAGAGGCTCAACCTGACAG 680
QY 603 CGTTGAAACTCAAAAGCTAGAGTGTGGCAGAGGAGTGAATTCATGTTAGCAGTG 662
DB 681 CGGTGAAACTCAAAAGCTAGAGTGTGGCAGAGGAGTGAATTCATGTTAGCAGTG 740
QY 663 AATGCGTGAAGATTAAGAAAGCAATGATGCGAAGAGGAGCTCTGAGGTTAACAATA 722
DB 741 AATGCGTGAAGATTAAGAAAGCAATGATGCGAAGAGGAGCTCTGAGGTTAACAATA 800
QY 723 CGCTATGACAGAAAGCGTGGGAGGACAAACAGATTATGATACCTGG--TAGTCAAGCC 780
DB 801 CACTGAGACAGAAAGCTAGAGGAGGACAAATGGAATTAGAGACCCAGATTAATCATAGCC 860
QY 781 CTAAACGATGTCATGATGTTGTTGGGCTTATTAAGCTTG--GTAAAGAACTAAGCGCT 838
DB 861 GTAAACGATGTCATGATGTTGTTGGGCTTATTAAGCTTG--GTAAAGAACTAAGCGCT 920
QY 839 GAAGTTGACCGGCTGGGAGTACCGTCCGAAAGTTAAACTCAAGAAATGACGGGAGC 898
DB 921 TAAGTATCCGCTGGGAGTACCGTCCGAAAGTTAAACTCAAGAAATGACGGGAGC 980
QY 899 CCGCAAGAGGAGTGAATTAAGTGAATTAATTCGATCAACGCGAAACCTTACCTACC 958
DB 981 CCGCAAGAGGAGTGAATTAAGTGAATTAATTCGATCAACGCGAAACCTTACCTACC 1040
QY 959 TTGACATGATGCGAATTTTCTAGAGATGATTAAGTG-CTTGGGAAACGCTAACAAGTG 1017
DB 1041 TTGACATGCGCGAATTTTCTAGAGATGATTAAGTG-CTTGGGAAACGCTAACAAGTG 1100
QY 1018 CTGATGAGCTGTGCTGACGCTCGTGTGATGATGTTGGGTTAGTCCGCAAGAGCGCA 1077
DB 1101 GTGCAATGAGCTGTGCTGACGCTCGTGTGATGATGTTGGGTTAGTCCGCAAGAGCGCA 1160
QY 1078 ACCCTGTCAATTAATGCTCAATCTTGTGGGCACTTAAATGAGACTGCGGTGACAA 1137
DB 1161 ACCCTGTCAATTAATGCTCAATCTTGTGGGCACTTAAATGAGACTGCGGTGACAA 1219
QY 1138 CCGAGAGAGGAGTGGAGTGAATGCTCAATGCTCAATGCGCTTAATGAGGCTTCAACAG 1197
DB 1220 CCGAGAGAGGAGTGGAGTGAATGCTCAATGCTCAATGCGCTTAATGAGGCTTCAACAG 1279
QY 1198 TAATCAATGCGCGCTAACAAGGCTTCCCAACCGGAGGAGGAGCTTAATCTGAAAGC 1257
DB 1280 TGCTCAATGAGGCGGAGCAAAAGGCTGCGATTCGAGAGGAGTGAATCTCAAAAAC 1339
QY 1258 GCGTGTGATGCGGAGTGGAGTCTGCAACTGCACTCGTGAATGCGAATCGCTAGTAAT 1317
DB 1340 GCTCTCAATGCGGAGTGGAGTCTGCAACTGCACTCGTGAATGCGAATCGCTAGTAAT 1399
QY 1318 CCGGATCAAG-CAATGCGGAGTGAATGCTTCCGAGCTTGTGAACACACCGCCGCTACA 1376
DB 1400 CGCGGATCAAG-CAATGCGGAGTGAATGCTTCCGAGCTTGTGAACACACCGCCGCTACA 1459
QY 1377 CCATGGAAGTGGCTTCAACAAGAGAGTGAATG-TCACGTTAAGAGAGGCGCTGAC 1435
DB 1460 CTATAGAGAGTGGCGATGTTGAATGATTAATCACTTAACGTTAAGAGAGGAGTGAAT 1519
QY 1436 GGTGAGATTCATGACTGGGGTG 1457
DB 1520 GGTGAGATTCATGACTGGGGTG 1541
```

RESULT 3  
BH814966/c 1143 bp DNA linear GSS 31-MAY-2003  
LOCUS BH814966  
DEFINITION pBPC15 Zea B chromosome PCR DNA library Zea mays genomic clone  
ACCESSION BH814966  
VERSION BH814966.1 GI:31249923  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACMAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 1143)

Cheng, Y.M. and Lin, B.Y.  
Cloning and characterization of maize B chromosome sequences derived from microdissection

Genetics 164 (1), 299-310 (2003)

12750341

COMMENT  
Contact: Bor-yaw Lin  
Department of Molecular Biology  
National Chung Hsing University  
250 Kuo Kuang Rd., Taichung 402, Taiwan (ROC)  
Tel: 886-4-2285-1885  
Fax: 886-4-2287-4879  
Email: bylin@dragon.nchu.edu.tw  
Insert Length: 1143 Std Error: 0.00  
Class: PCR fragment.

FEATURES  
source  
location/Qualifiers  
1..1143  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="1289"  
/db\_xref="taxon:4577"  
/clone="pBPC15"  
/sex="hermaphrodite"  
/cell\_type="Pollen mother cells"  
/dev\_stage="Meiosis I"  
/lab\_host="Drosophila"  
/clone\_lib="Maize B chromosome PCR DNA library"  
/note="Vector: pBSK-; The library was constructed from microdissected B chromosome DNA."

ORIGIN  
Query Match 44.1%; Score 648.4; DB 9; Length 1143;  
Best Local Similarity 89.6%; Pred. No. 1.5e-188;  
Matches 744; Conservative 0; Mismatches 76; Indels 10; Gaps 4;

631 CAGAGGAGGTGGAATTCATGCTGTGTAGCACTGAATTCGCTGAGATATGGAAGAATCG 690  
|||||  
1143 CAGAGGGGGGTAAATTCACGCTGTAGCACTGAATTCGCTGATATGTGAGAGAACACCG 1084  
|||||  
691 ATGGCGAAGGACGCTCTGGGTTAACAAGAGCTCATGACAGAAAGCGTGGGAGCAA 750  
|||||  
1083 ATGGCGAAGGACGCTCTGGGTTAACAAGAGCTCATGACAGAAAGCGTGGGAGCAA 1024  
|||||  
751 ACAGATTAGATACCTGTAGTCAAGCCCTTAACAGATGTCACTAGTTGGGCTT 810  
|||||  
1023 ACAGATTAGATACCTGTAGTCAAGCCCTTAACAGATGTCTAGTTGGGCTT 964  
|||||  
811 ATTAGCTGTGTAACGATTAACGCTGAAGTTACCGCTGGGAGTACGGTCCGCAAG 870  
|||||  
963 AATTGACGTTGTAAACGATTAACGCTGAAGTTACCGCTGGGAGTACGGTCCGCAAG 904  
|||||  
871 ATTAAGCTCAAGGAATTGAGCGGGGACCGGCAAGCGGTGATTAATGATTAATTC 930  
|||||  
903 ATTAAGCTCAAGGAATTGAGCGGGGACCGGCAAGCGGTGATTAATGATTAATTC 844  
|||||  
931 GATGCAACGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 990  
|||||  
843 GATGCAACGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 784  
|||||  
991 AGTGCT--TCGGGAACGCTTAACAGAGTCTGATGCTGTGTGTGTGTGTGTGTGTGTGA 1047  
|||||  
783 AGTGCTGGAAGGAACAGTACAGAGTCTGATGCTGTGTGTGTGTGTGTGTGTGTGTGA 724  
|||||  
1048 GATGTTGGGTTAGTCCCGCAAGAGGCAACCTTGTCAATTAATGCGCATTTGGTT 1107  
|||||  
723 GATGTT--GGTTAGTCCCGCAAGAGGCAACCTTGTCAATTAATGCGCATTTGGTT 670  
|||||  
1108 GGGCACTTAATGAGCTGCGGTGACAAACCGAGAGAAAGTGGGAGTGAAGTCAAGTCC 1167  
|||||  
669 GGGCACTTAATGAGCTGCGGTGACAAACCGAGAGAAAGTGGGAGT--ACGTCAAGTCC 611

1168 TCATGCGCCCTTATGGGTAGGGCTTCACAGTAAATCAATGGCGGTACAGAGGGTTGCCA 1227  
|||||  
610 TCATGCGCCCTTATGGGTAGGGCTTCACAGTAAATCAATGGGTACAGAGGGTTGCCA 551  
|||||  
1228 ACCCGGAGGGGAGCTAATCTCAGAAAGCGCGTGTAGTCCGATCGGAGCTTGCAACT 1287  
|||||  
550 ACCCGGAGGGGAGCTAATCTCAGAAAGCGCGTGTAGTCCGATCGGAGTTAGCTGCAACT 491  
|||||  
1288 CGACTTCGTTGAAGTGTGAATGCTTGTATGCGCGATTCAGCATGTCTGGGTGAATACGTT 1347  
|||||  
490 CGACTTCGTTGAAGTGTGAATGCTTGTATGCGCGATTCAGCATGTCTGGGTGAATACGTT 431  
|||||  
1348 CCGGGGTCTTGTACACACCGCCCGTACACCATGAGGAGTGGGTTTACACAGAACGATTA 1407  
|||||  
430 CCGGGGTCTTGTACACACCGCCCGTACACCATGAGGAGTGGGTTTACACAGAACGATTA 371  
|||||  
1408 GTCTAACCGTAAAGAGGCGCTTGCACAGGTGAGATTGATGACTGGGGTG 1457  
|||||  
370 GCTTAACCGCAAGAGGCGGCTTACACAGGTGAGATTGATGACTGGGGTG 321  
|||||

RESULT 4  
BZ439740 1084 bp DNA linear GSS 13-DEC-2002  
LOCUS BZ439740  
DEFINITION BONDUS4TF BO.1.6.2 KB tot Brassica oleracea genomic clone BONDUS4,  
genomic survey sequence.  
ACCESSION BZ439740  
VERSION BZ439740.1 GI:26695676  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosids; eustroids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 1084)  
Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utecherback, T.R., Mortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490  
Other GSSs: BONDUS4TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.

FEATURES  
source  
location/Qualifiers  
1..1084  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONDUS4"  
/clone\_lib="BO.1.6.2 KB tot"  
/note="Vector: pHOS1; Site 1: BactXI; 1.6-2 kb sheared total DNA inserted into pHOS1 using BactXI linkers"

ORIGIN  
Query Match 44.1%; Score 643; DB 9; Length 1084;  
Best Local Similarity 77.3%; Pred. No. 7.1e-187;  
Matches 842; Conservative 0; Mismatches 220; Indels 27; Gaps 4;

237 CTACCAAGCGCAGTACGTAGTGTGCTGAGAGGACGACGACGACGACGACGACGACGACG 296  
|||||  
DB 3 CTACCAAGCGCAGTACGTAGTGTGCTGAGAGGACGACGACGACGACGACGACGACGACG 61  
|||||  
QY 297 CACGCCCAAGCTCTTAACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 356  
|||||

Db 62 CAGGCCCCAGACTCTTACGGAGAGCAGCACTGAGGAAATTTTCCGCAATGAGGCGAAAGCCCT 121  
 Qy 357 GATCCAGCAATGCGCGAGTAGTGAAGAGGCTTTCGGGTTGTAAGCTCTTTCAGTCCAG 416  
 Db 122 GACGAGCAATGCGCGAGTAGTGAAGAGGCTTTCGGGTTGTAAGCTCTTTCAGTCCAG 181  
 Qy 417 AAGAAAGGTTACGGTAAATATCTGTGACCCATGACCGTATCGACAGAGAGACCCGCGC 476  
 Db 182 AAGA-----AGCAATGACGGTATCTGGGGAAATTAACATCGCGC 218  
 Qy 477 TAACTAGTCCGACAGACCCCGGTAATACGTAGGGTGCAGACCTTATCGGAATTTCTG 536  
 Db 219 TAACTGTCTCCAGCAGCCCGGTAATACAGAGATGCAACCGTTATCCGGAATGATGG 278  
 Qy 537 GCGTAAAGGGTGGCAGCGCGCTTGTAGTGCAGATGTGAATCCCGGGCTTAACTCGG 596  
 Db 279 GCGTAAAGGGTGTGTAGTGGCTTTTAAAGTCCGCGCTCAATCCGAGGCTCAACCTTG 338  
 Qy 597 GAAATGCGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTA 656  
 Db 339 GACAGGGGAGTGAAGAACTAACAGCTTGAAGTGGGAGAGGGAATTTCCGGTGGGA 398  
 Qy 657 GCACTGAAATGCGTAGAGATATGGAAGAAATCGATGCGGAAGCAGCTCTTGGGTTAA 716  
 Db 399 GCGGTGAATGCGTAGAGATCGGAAGAAACCAACGCGCAAGCACTCTGCTGGGCGGA 458  
 Qy 717 CACTAGAGCTCATGACAGAAAGCGTGGGGAGCAAGAGATTAGATACCTCGGTAGTCCA 776  
 Db 459 CACTAGACATGAGAGAGAAAGCTAGGGAGCGAATGGATTAATATCCCAAGTAGTCT 518  
 Qy 777 CGCCCTAAAGGATGTCAATAGTTGTTGGGCTTATTAAGCTTG--GTAAAGAGCTTAC 834  
 Db 519 AGCCGTAAGAGATGAGATCTAGAGGCGTGTGGTATCAACCGGTGCAAGTGTAGTAC 578  
 Qy 835 GCGTGAAGTGAACCGCTGGGAGTACGGTGCAGATTAATAACTCAAGAAATTTAGCG 894  
 Db 579 GCGTTAAGTATCCCGCTGGGAGTACGTTGCAAGATGAAATCTCAAGAGAAATTTAGCG 638  
 Qy 895 GAGCCCGCAAGGAGGTGAGTATGATGATTAATTCAGTCAAGCGCAAGAACTTACTCT 954  
 Db 639 GGGCCCGCAAGGAGGTGAGTATGATGATTAATTCAGTCAAGCGCAAGAACTTACTCA 698  
 Qy 955 ACCCTTGAATGTAGCGAATTTTCTAGAGATGATTAAGTGTG--CTTGGGAAAGCTTACACA 1013  
 Db 699 GGGCTTGAATGTAGCGAATTTTCTAGAGATGATTAAGTGTG--CTTGGGAAAGCTTACACA 758  
 Qy 1014 GGTGCTGAGTGGCTGTGCTGAGCTGTGTGCTGAGATGTTGGTTAAGTCCCGAAGAG 1073  
 Db 759 GGTGCTGAGTGGCTGTGCTGAGCTGTGTGCTGAGATGTTGGTTAAGTCCCGAAGAG 818  
 Qy 1074 CGCAACCTTGTCAATTAATTTGCCATCATTTTGGTGGGCACTTATATGAGATCGCGGGA 1133  
 Db 819 CGCAACCTTGTCAATTAATTTGCCATCATTTTGGTGGGCACTTATATGAGATCGCGGGA 878  
 Qy 1134 CAAACCGGAGAGGTGGGAGTGAAGTCAAGTCCCTCATGGCTTATGGTAGGAGCTTCA 1193  
 Db 879 TAAGCCGAGAGAGGTGAGATGAGCTCAATCATATGCCCTTATGCCCTGCGGAGCA 938  
 Qy 1194 CACGTAAATACATGCGCGGTAAGAAGGTTGCCAACCCGCGAGGGGAGGATTAATCTCA 1253  
 Db 939 CACGTGCTACATGCGCGGTAAGAAGGTTGCCAACCCGCGAGGGGAGGATTAATCTCA 998  
 Qy 1254 AAGCGCTGTAGTCCGATCGGAGTCTGCAATCTCGATCCGTGAAGTCCGAATCGCTAG 1313  
 Db 999 AAGCGCTGTAGTCCGATCGGAGTCTGCAATCTCGATCCGTGAAGTCCGAATCGCTAG 1058  
 Qy 1314 TAATCGCGG 1322  
 Db 1059 TAATGCGG 1067

RESULT 5

CK296497  
 LOCUS CK296497 941 bp mRNA linear EST 02-AUG-2004  
 DEFINITION EST759211 Nicotiana benthamiana mixed tissue cDNA library.  
 normalized, full-length Nicotiana benthamiana cDNA clone NBMD226 5' end, mRNA sequence.  
 CK296497  
 ACCESSION CK296497.1 GI:39881943  
 VERSION  
 KEYWORDS  
 SOURCE Nicotiana benthamiana  
 ORGANISM Nicotiana benthamiana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Solanales; Solanaceae; Nicotiana.  
 REFERENCE 1 (bases 1 to 941)  
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B., Staskiewicz, B., Jin, H. and Baker, B.  
 Generation of EST sequences from Nicotiana benthamiana  
 Unpublished (2003)  
 TITLE  
 JOURNAL  
 COMMENT Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.  
 Seq primer: APT TNG CTC ACA CTA TNG.  
 FEATURES  
 source  
 1..941  
 /organism="Nicotiana benthamiana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4100"  
 /clone="NBMD226"  
 /tissue\_type="adictic and biotic stress-treated leaves, callus tissue and root tissue"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"  
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr; and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."  
 ORIGIN  
 Query Match 43.7%; Score 636.2; DB 7; Length 941;  
 Best Local Similarity 81.2%; Pred. No. 8.7e-185;  
 Matches 751; Conservative 0; Mismatches 173; Indels 1; Gaps 1;  
 Qy 1 ATTGAAGCTGGCGGATGTTTACATCAAGTGGAAAGGAGCAGCGATGCTTGCAT 60  
 Db 18 ATTGAAGCTGGCGGAGGCTTAACATGACATGAGCGGAGCAAGGCTACTTGTAC 77  
 Qy 61 CTGGTGCAGTGGCGGAGCGGTGATGATCGGACGTATCCAGAAAGAGGGGGTA 120  
 Db 78 CTGGTGCAGGCGGCGGAGCGGTGATGATGCTTAGGAATCTGCTGTGTGGGGGATA 137  
 Qy 121 ACGCATGAAGAATGTGCTAATACCGCATATCTTATAGAGAGAAAGAGGGATCGAAA 180  
 Db 138 ACGCTCGGAAGCGAACCTTAATACCGCATATCGTCTACGGAGAAAGACAGGGGACTTCG 197  
 Qy 181 GACCTTGGCTTTTGGAGCGCGCGATGCTGATTAGTATGGTGGGTAAAGGCTTAC 240  
 Db 198 GGCCTTGGCTTTTGGAGCGCGCGATGCTGATTAGTATGGTGGGTAAAGGCTTAC 257  
 Qy 241 CAAAGCGAGATCAGTATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 Db 258 CAAAGCGAGATCAGTATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 317



Qy	1037	TCGTCGTGTCGATGTTGGGTTAGTCTCCGGAAAGAGCGCAACCTTGTCACTTAATTTAGCC	1096
Db	236	TCGTGCGCTTAAGTGTGTGGTTAACTCCCGCAAGAGCGCAACCTGTGTTTTAGTTAGTCC	177
Qy	1097	ATCATTTTGGTTGGGCACTTTAATGAGACTGCGCGGTGACAAACCGAGAGAGTGGGGATG	1156
Db	176	ACCGTTGAGTTTGGAAACCTTGAAACGACTGCGCGGTGATTAACCGGAGGAAAGTGAAGATG	117
Qy	1157	ACGTCGAATGCTCATGCGCCTTTATGGGTAAAGGGCTTCACAGCTAATTAACAATGGCCGCTACA	1216
Db	116	ACGTCGAATGCTCAAGCCCTTTATGCGCCTGGAGCACAGCTGTCTACATATGACCGGGACA	57
Qy	1217	GAGGGTTCACCAACCCGCGAGGGGAGACTAATCTCAGAAAGCGCGTGTAGTCCGGA	1272
Db	56	AAGGCTCGCGATCCAGAGAGGGGTAGCTAATCTCCAAAACCCGTCTCAGTTCCGA	1
RESULT 7			
LOCUS	BZ431523/c	1010 bp	DNA
DEFINITION	BONHD88TF_BO.1.6.2_KB_tot	Brassica oleracea	linear GSS 13-DEC-2002
ACCESSION	BZ431523		
VERSION	BZ431523.1		
KEYWORDS	GI:26677909		
SOURCE	GSS.		
ORGANISM	Brassica oleracea		
REFERENCE	Brassica oleracea		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.		
TITLE	1 (bases 1 to 1010)		
JOURNAL	Ayele, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,		
PUBMED	Uiterbeek, T. R., Mortman, J. R., White, O. R. and Town, C. D.		
COMMENT	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res. 15 (4), 487-495 (2005)		
OTHER GSSs:	BONHD88TR		
CONTACT:	Chris Town		
TIGR	9712 Medical Center Drive, Rockville, MD 20850, USA.		
TELEPHONE:	301-838-3523		
FAX:	301-838-0208		
EMAIL:	cdtown@tigr.org		
NOTE:	DNA is from a doubled haploid provided by Tom Osborn.		
SEG PRIMER:	TF		
CLASSES:	sheared ends.		
LOCATION/QUALIFIERS	1..1010		
ORGANISM:	"Brassica oleracea"		
MOLECULAR TYPE:	"genomic DNA"		
STRAIN:	"TO1000DH3"		
DB XREF:	"taxon:3712"		
CLONE:	"BONHD88"		
CLONE 11b:	"BO.1.6.2_KB_tot"		
NOTE:	"Vector: pHD1; Site 1: BstXI; 1.6-2 kb sheared total DNA inserted into pHD1 using BstXI linkers"		
ORIGIN			
Query Match	43.2%;	Score 629.4;	DB 9; Length 1010;
Best Local Similarity	78.0%;	Pred. No. 1.1e-182;	
Matches	806;	Conservative 0;	Mismatches 181; Indels 26; Gaps 3
Db	211	GATTACTGTTGGTGGGGTAAAGGCTACCAAGGAGCAATCAGTATGGTCTGAGAG	270
Qy	1010	GATTACTGTTGGGTAGGCAATAGCTTACCAAGGCAATCAGTATGGTCTGAGAG	951
Db	271	GAAGACCAACCACTGAGACTGAGACACAGGCGCCAGACTCTTACGGGAGCAGCAGTGGG	330
Qy	950	GATGATCAACCACTGAGACTGAGACACGCGCCAGACTCTTACGGGAGCAGCAGTGGG	891
Db	331	GAATTTTGGACATGGGCGCAACCTGATTCAGCAATGCCGGGTAGTAAAGAGCCCTT	390

[illegible]

ORGANISM *Brassica oleracea*  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Utecherback,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
TITLE Whole genome shotgun sequencing of *Brassica oleracea* and its  
application to gene discovery and annotation in Arabidopsis  
JOURNAL Genome Res. 15 (4), 487-495 (2005)  
PUBMED 15805490  
COMMENT Other GSSs: BOMKA47TR  
Contact: Chris Town  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TP  
Class: sheared ends.  
FEATURES  
source Location/Qualifiers  
1..1079  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone\_lib="BO\_2\_3\_KB"  
/note="Vector: pHD1; Site 1: BstXI; 2-3 kb sheared  
genomic DNA inserted into pHD1 using BstXI linkers"  
ORIGIN  
Query Match 42.8%; Score 623.8; DB 9; Length 1079;  
Best Local Similarity 76.6%; Pred. No. 6.2e-181;  
Matches 813; Conservative 0; Mismatches 222; Indels 26; Gaps 3;  
186 TGGCGTTTGGAGCGGCCCATGTCGATTAGCTAGTGTGGGGGTAAGGCTTAAAGC 245  
42 TCGGCCGAGAGAGGGCTCGCGTCTATTAGTGTGGAGGCAATAGCTTACCAAG 101  
246 CGACGATCAGTAGTGTGCTGAGAGAGCAGCAGCACACTGGAGCTGAGACAGCGCCCA 305  
102 CGATGATCAGTAGCTGTGCGAGAGATATCAAGCAGCTGGAGCTGAGACAGCGCCCA 161  
306 GACTCTTACGAGAGCAGCAGTGGGAATTTGACATATGGCGCAAGCTGATCCAGCA 365  
162 GACTCTTACGAGAGCAGCAGTGGGAATTTGCGCAATGGCGCAAGCTGACGAGCA 221  
366 ATGCCCGGAGAGAGAGGCTTGGGGTGTAAAGCTCTTCAATCGAGAGAGAGAG 425  
222 ATGCCCGGAGAGAGGCTTACGCGGCTTGAACCTTTTCCAGAGAGAGAGAGAG 276  
426 TTACGGTAATAATCGTGACCCATGACGATGACAGAGAGAGAGAGAGAGAGAG 485  
277 -----AGCATGACGATATCTGGGGAATAAGCATCGGCTAACCTCTGT 318  
486 GCCAGAGCGCGGTAAATAGTAGGAGTGAAGCGTTAATCGAATTAATGAGCGTAAAG 545  
319 GCCAGAGCGCGGTAAATAGTAGGAGTGAAGCGTTAATCGAATTAATGAGCGTAAAG 378  
546 GTGCGAGCGCGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 605  
379 GTCTGAGAGTGGTTTAAAGTCCGCGTCAATTCAGAGGCTCAACCTTGAAGAGCG 438  
606 TTGAAGCTCAAAAGCTAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 665  
439 TGGAAACTACAAAGCTTGAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 498  
666 TGGTAGAGATGAG 725  
499 TGGTAGAGATGAG 558

QY 726 TCATGACAGAAAGCGTGGGAGCAAAAGAGATTAGATACCTGTAGTCCAGCCCTTAA 785  
DB 559 TGAAGAGAGAAAGCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 618  
QY 786 CGATGCACTAGTGTGGGCTTATTAAGCTTG--GTACGAAGCTTACCGCTGAAGT 843  
DB 619 CATGAGATCTAGGGGCTGTGTGTGATGCAACCCGTGACAGTGTGTGCTAAACCGCTTAA 678  
QY 844 TCACCGCGTGGGAGATAGAGTGGCAAGATTAAGTCAAAAGAGATGACGAGGAGCCGCA 903  
DB 679 ATCCGCGTGGGAGATAGAGTGGCAAGATTAAGTCAAAAGAGATGACGAGGAGCCGCA 738  
QY 904 CAAGCGTGGATTAATGAGATTAATTCATGACAGCGCAAAAGCTTACCTTACCTTAC 963  
DB 739 CAAGCGTGGAGATAGTGTGTTAATTCATGACAGCGCAAAAGCTTACCTTACCTTAC 798  
QY 964 ATGAGCGAATTTCTAGAGATAGATTAAGT--CTTGGGAGAGCTTACAGAGTGTGCA 1022  
DB 799 ATGCCGCAATCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 858  
QY 1023 TGGCTGTGTGACGCTCGTGTGTGAGATGTTAGTCCCGCAAGAGGCAACCTT 1082  
DB 859 TGGCTGTGTGACGCTCGTGTGTGAGATGTTAGTCCCGCAAGAGGCAACCTT 918  
QY 1083 TGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1142  
DB 919 CGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 978  
QY 1143 GGAAGGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1202  
DB 979 GGAAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1038  
QY 1203 CAATGCGCGTACAGAGGTTGCCAAGCCGAGAGGAGGAGG 1243  
DB 1039 TCATGCGCGGAGCAAAAGGTCGATCCCGAGAGGAGTGAAGC 1079  
RESULT 9  
BZ426201/C 1016 bp DNA linear GSS 13-DEC-2002  
LOCUS BOND228TF BO.1.6.2\_KB\_tot Brassica oleracea genomic clone BOND228,  
DEFINITION genomic survey sequence.  
ACCESSION BZ426201  
VERSION BZ426201.1 GI:26667713  
KEYWORDS GSS  
SOURCE  
ORGANISM *Brassica oleracea*  
*Brassica oleracea*  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 1016)  
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Utecherback,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
TITLE Whole genome shotgun sequencing of *Brassica oleracea* and its  
application to gene discovery and annotation in Arabidopsis  
JOURNAL Genome Res. 15 (4), 487-495 (2005)  
PUBMED 15805490  
COMMENT Other GSSs: BOND228TR  
Contact: Chris Town  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TP  
Class: sheared ends.  
FEATURES  
source Location/Qualifiers  
1..1016  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"

## ORIGIN

/clone="BOND28"  
/clone\_1ib="BO.1.6.2\_KB\_tot"  
/note="Vector: PHOS1; Site 1: BexXI; 1.6-2 kb sheared  
total DNA inserted into PHOS1 using BexXI linkers"

Query Match 42.8%; Score 623.4; DB 9; Length 1016;  
Best Local Similarity 78.1%; Pred. No. 8.1e-181;  
Matches 811; Conservative 0; Mismatches 201; Indels 27; Gaps 4;

QY 200 GGCCTATGCTCTATTTAGCTTATTTGTTGGGCTAAAGGCTTACCAAGGCGAGTCACTACT 259  
DB 1015 GGCCTGCGCTCTATTTAGCTTATTTGTTGGGCTAAAGGCTTACCAAGGCGAGTCACTACT 956  
QY 260 TGGTCTGAGAGACGACGACCACTGAGGAGTGAAGACAGCGGCCAGACTCTTACGCGAG 319  
DB 955 TGGTCTGAGAGATGATTCAGCCACTGGGAGTGAAGACAGCGGCCAGACTCTTACGCGAG 896  
QY 320 GAGCAGTGGGGAATTTGAGCAATGGGCGCAAGCTTATCCAGCAATGCCGCTGAGTG 379  
DB 895 GAGCAGTGGGGAATTTGCGG-CAATGGGCGAAAGCTTACGAGCAATGCCGCTGAGTG 837  
QY 380 AAGAGGCTTGGGCTTGTAAAGCTTTCACTGCGAGGAAGAAAGTTACGTAATTAAT 439  
DB 836 TAGAAGGCTTACGCGCTTGAACCTTTTCCAGAGAAAG----- 796  
QY 440 CGTGAACCATGATGATGATGACAGAGAAAGACCGGCTAATCTGTCGACGACGCGCG 499  
DB 795 ----ACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 740  
QY 500 TAATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 559  
DB 739 TAATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 680  
QY 560 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 619  
DB 679 TTTTAAAGTCCGCGCTGAAATCCAGAGGCTTACCTGAGAGGCGGCTGAAATCTACCA 620  
QY 620 CTAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 679  
DB 619 CTAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 560  
QY 680 GAAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 739  
DB 559 AAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 500  
QY 740 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 799  
DB 499 TAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 440  
QY 800 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 857  
DB 439 CGTGTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 380  
QY 858 GTACGCTGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 917  
DB 379 GTACGCTGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 320  
QY 918 TGTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 977  
DB 319 TGTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 260  
QY 978 CTAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1036  
DB 259 CTAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 200  
QY 1037 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1096  
DB 199 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 140  
QY 1097 ATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1156  
DB 139 ACCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 80

QY 1157 AGCTAAGTCTCTATTTAGCTTATTTGTTGGGCTAAAGGCTTACCAAGGCGAGTCACTACT 1216  
DB 79 AGCTAAGTCTCTATTTAGCTTATTTGTTGGGCTAAAGGCTTACCAAGGCGAGTCACTACT 20  
QY 1217 GAGGCTGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1235  
DB 19 AAGGCTGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1

## RESULT 10

BH656222/c 1031 bp DNA linear GSS 19-FEB-2002

LOCUS BOMZ86TF BO.2\_3\_KB Brassica oleracea genomic clone BOMZ86, 1031 bp

DEFINITION genomic survey sequence.

ACCESSION BH656222

VERSION BH656222.1 GI:18714532

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 1031)

Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utecher, T.R., Wortman, J.R., White, O.R. and Town, C.D.

Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis

Genome Res. 15 (4), 487-495 (2005)

15805490

COMMENT Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: 1F

Classes: sheared ends.

Location/Qualifiers

1..1031

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="TO1000DH3"

/db\_xref="taxon:3712"

/clone="BOMZ86"

/clone\_1ib="BO.2.3\_KB"

/note="Vector: PHOS1; Site 1: BexXI; 2-3 kb sheared

genomic DNA inserted into PHOS1 using BexXI linkers"

## ORIGIN

Query Match 42.5%; Score 619; DB 9; Length 1031;

Best Local Similarity 76.6%; Pred. No. 1.9e-179;

Matches 807; Conservative 0; Mismatches 220; Indels 26; Gaps 3;

QY 307 ACTCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 366  
DB 1031 ACTCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 972  
QY 367 TGGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426  
DB 971 TGGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 918  
QY 427 TACGTAATTAATCGTACCATGACGATGACGATGACGATGACGATGACGATGACGATG 486  
DB 917 -----ACCATGACGATGACGATGACGATGACGATGACGATGACGATGACGATG 875  
QY 487 CCAGCAGCGCGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 546  
DB 874 CCAGCAGCGCGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 815  
QY 547 TGGCAGCGCGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 606



DB	81.4	TCGTAGTGGCTTTTAAAGTCGGCCGTCAATCCAGGGCTCAACCTTGACAGGGCGT	75
QY	607	TGAACCTACAAAGCTTAGAGTGTGGCAGAGGGAGGTGAAATTCATGTGTAGCATGTAAAT	666
Db	754	GGAAACTACCAAGCTTGAAGTACGGTAGGGGCAGAGGGAAATTTCCGGTGAAGCGGTGAAT	695
QY	667	GGGTAGAGTATAGGAAGAACATCGAATGGGAAAGGCGAGCTCCGGGGTTAAACATGACGT	726
QY	727	CATGCACGAAACGCTGGGAGGCAACAGATTAGATACCTGGTAGTCCAGCCCTTAAC	786
Db	634	GAGAGACGAAACCTAGGGGAGGGAATGGGATTAGATACCAGTAGTAGTCTACCGCTTAAC	575
Db	694	GGTAGAGATCGGAAGAACACCAACGGGAAAGCATCTGTCTGGGCCGACATGACACT	635
QY	787	GATGTCAACTAGTGTGGGCCCTTATTAGGCTTG--GTACGAAAGCTTAACGGTGAAGTT	844
Db	574	GATGATATCTAGCGCCTGTGCGTATCGAACCCGTGCAGTCTGTAGCTTAACGGGTTAAGTA	515
QY	845	GACCCCTCGGGGAGTAACGGTGGCAAGTTAAACCTAAAGGAATTGACGGGGACCCGGAC	904
Db	514	TCCCCCTCGGGGAGTACGTTGCGAAGAAAGAACTCAAGGAATGACGGGGGCCGGAC	455
QY	905	AAGCGGTGATATGTGATTTAATTCGATGCACAACGGAAAAAAGCTTACCTACCTTGAACA	964
Db	454	AAGCGGTGAGCATGTGTTTAATTCGATGCACAAGGAAGAAAGCTTACCAAGGGCTTGAACA	395
QY	965	TGTACGGAATTTTCTAGAGATAGATTATGTC-CTTCGGGAACGCTTAACAGAGTGTGCAT	1022
Db	394	TGCCCGGAATCTCTTGAAGAGAGGGGTGCTTCGGGAACCGGACACAGGTGTGTGCTAT	335
QY	1024	GGCTGTCTGACGCTGTGTCTCGGAGATGTGTGGGTTAATGTCCCGCAACGAGCCCAACCTT	1082
Db	334	GGCTGTCTGACGCTGTGTCTCGGAGATGTGTGGGTTAATGTCCCGCAACGAGCCCAACCTTC	275
QY	1084	GTCAATTAATGTGCATCATTTGTGGGCACTTTAATGAGACTGCGGTGACAAACCGGAG	1142
Db	274	GTGTTTAGTGTGCCACCGTTGAATTTGGAACCTTGACAGACATGCGGTGATAGCCGGAG	215
QY	1144	GAAAGGTGGAGTGAACGTCAAGTCTCATAGGCCCTTAATGAGGCTTACACAGTAAATAC	1202
Db	214	GAAAGGTGAGATGACGTCAAGTCAATCATATCCCTTAATGCGGGGCAACACAGTGTATAC	155
QY	1204	AATGGCGGTGACAGAGGGTGTGCCAACCCGCGAGGGGGAGACTAATCTCAGAAAGCGGCTG	1262
Db	154	AATGGCGCGGACAAAGGGGTGCGGATCCCGCGAGGGTGACCTAACCTCAAAAACCCGTCT	95
QY	1264	TAGTCGGATGCGAGCTGCAACTGCACTCCGATGAAGTGGGAATGCGTATAGCGGGA	1322
Db	94	CAGTTTCGATTGCAAGCTGCAACTGACTGATGAAGCCGGAATGATATGATATGCGCG	35
QY	1324	TCAGCATGTCCGCGTGAATACGTTCCCGGGTCT	1356
Db	34	TCAGCCATATCGGCGGGAATTCGTTCCCGGGCCT	2
RESULT 11			
LOCUS	BZ450751	1077 bp	DNA
DEFINITION	BOMB063TR.B0.1.6.2_KB	loc	Brassica oleracea genomic clone BOMB063,
ACCESSION	BZ450751		genomic survey sequence.
VERSION	BZ450751.1	GI:26721334	
KEYWORDS	GSS.		
SOURCE	Brassica oleracea		
ORGANISM	Brassica oleracea		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;		
	rosids; eurosids II; Brassicales; Brassicaceae; Brassica.		
	1 (bases 1 to 1077)		
	Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,		
	Utecherback, T.R., Wortman, J.R., White, O.R. and Town, C.D.		
TITLE	Whole genome shotgun sequencing of Brassica oleracea and its		
	application to gene discovery and annotation in Arabidopsis		

JOURNAL	Genome Res. 15 (4), 487-495 (2005)
PUBMED	15805490
COMMENT	Other_GSSS: BONB063TF Contact: Chris Town TIGR
	9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdclown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TR Class: sheared ends.
FEATURES	Location/Qualifiers
source	1..1077 /organism="Brassica oleracea" /mol_type="genomic DNA" /strain="TO1000DH3" /db_xref="taxon:3712" /clone="BONB063" /clone_jid="BO_1.6.2_KB_tor" /note="Vector: pHD1; Site:1; BstXI, 1.6-2 kb sheared total DNA inserted into pHD1 using BstXI linkers"
ORIGIN	
Query Match	42.3%; Score 616; DB 9; Length 1077;
Beet Local Similarity	77.4%; Pred. No.1.6e-178; Mismatches 225; Indels 4; Gaps 3
Matches	785; Conservative 0;
444	ACCATGACGGTATCGACAGAAAGACCGGCTTACCTGTCGACAGCCGCGGTAT
64	AGCATATGACGGTATCGGGAAATAGATCGGCTTACCTGTGACAGCCGCGGTAT
504	ACGTAGGTCGACGGCTTAATCGAATTACTGGGCGTTAAGGGTGCAGCGCGCTTGT
124	ACAGAGATGCAAGGCTTATCCGGATGATTGGGCGTTAAACGCTGATGGTGGCTTTT
564	AAGTCAGATGGAATCCCGGGCTTACCTGGGAATTGCGTTGAACTTCAAAAGCTAG
184	AAGTCGCGCTCAATCCAGGCGCTCAACCTTGACAGCGCGGTGAACTTCAAGCTTG
624	AGTGTGACAGAGGAGTGGAATTCATGTGTAGCAGTGAATGCGTAGATATGAAAG
244	AGTACGATGAGGGCAGAGGAAATTTCCGCTGGAACGCGTAATGCTGATGAGATCGAAAG
684	AACATCATGCGCAAGGCGCTCTCGGTTTAACTGACGCTCATGCAAGAAACGCTGG
304	AACACCAACGCGCAAGGCACTCTGCTGGGCGCACTGACACTGAGACGAAAGCTAGG
744	GAGCGAAACAAGATTAGTAACTCCCTGTGTGTCAAGCCCTTAAACGATGCAACTAGTTT
364	GAGCGAATGGAATTAAGTATCCCACTGATGCTTCTTACCGTAAACGATGATCTAGGCGCT
804	GAGCCCTTATTAAGGCTTG--GTAAACGAAGCTAACGCGTGAATTTGACCGCTGGGAGTAC
424	GTGCGTATCGAACCGTGCAGTGTCTGTACTAACCGCTTAAATATTTCCGCTGGGAGTAC
862	GCTGCGAAGATTAATACTCAAGGAATTCAGCGGAGCCCGCAACAAGCGGTGATTATGTG
484	GTTGCGAAGATTAATACTCAAGGAATTCAGCGGAGCCCGCAACAAGCGGTGAGCATGTG
922	GATTATTCATGCAACCGCGAAACCTTACTTACCTTGTGACATGTAGACGAAATTTTCTAG
544	GTTTAAATTCATGCAACCGCGAAACCTTACTTACCGGCTTGTGACATGTAGACGAAATTTTCTAG
982	AGATGATTAATG--CTTGGGAGCGCTTAAACAGGTGTGATGAGTCTGTGCTAGCTGT
604	AAAGAGAGGGGTGCTTTCGGAAGCGCGGACACAGGTGTGATGAGTCTGTGCTAGCTGT
1041	GTCTGTGATGTTGGGTTAATGCTCCCGCAACAAGCGCAACCTTGTCAATTAATGGCATCA
664	GCGGTAAAGTGTGGGTTAATGCTCCCGCAACAAGCGCAACCTTGTGTAGTTGGCACCG
1101	TTTGGTTGGGCACTTTAATGACTGCCGCTGACAAACCGGAGGAAGGTGGGATGACGT



Db 724 TTGAGTTTGACCTTGACAGACTGCGGTGATTAAGCCGAGAGTGAAGATGACGT 783  
1161 CAAGTCTCATGGCCCTTATGGTAAAGGCTTCAACAGTAATACAAATGCGCGTACAGAG 1220  
Db 784 CAAGTCAATACGCCCCCTTATGCTTATGCGGCAACACAGTGTCAATAGCCGGGACAAAG 843  
Qy 1221 GTTGCCAAACCCGAGGAGGAGTAACTCTGAGAAAGCGCTGTAGTCCGATCGAGTC 1280  
Db 844 GTGCGCATCCCGAGAGGTGAGCTTAATCTCAAAAACCGTCTCAAGTTGACAGCG 903  
Qy 1281 TGCATCTCACTCCGTAAATCGGAAATCGCTAATACGCGGATCAG-CATGTGCGGCTG 1339  
Db 904 TGCATCTCCGCTCATTAACCCGAAATCGCTAATACGCGGATCAGCGGATCGGCTG 963  
Qy 1340 AATACGTTCCCGGCTCTTGTACACACCGCCGTACACCAATGGAGTGGTTTCCAGCA 1399  
Db 964 AATTCTGTTCCCGGCTCTTGTACACACCGCCGTACACCAATGGAGTGGGCAATGCCGA 1023  
Qy 1400 AGCAGTAGTCTTAACCGTAAGAGAGGCGCTTGCACGCTGATTCATGACTGG 1453  
Db 1024 AAGTCGTACCTTAACCGCAGAGAGGCGGCTGCCGAGAGCGCTAGTGACTGG 1077

RESULT 12  
BZ463550/c 1044 bp DNA linear GSS 13-DEC-2002  
LOCUS BONK86TR\_BO\_1.6.2\_KB\_tot Brassica oleracea genomic clone BONK86,  
DEFINITION genomic survey sequence.  
ACCESSION BZ463550  
VERSION BZ463550.1 GI:26749465  
KEYWORDS GSS.

## SOURCE

Brassica oleracea  
Brassicaceae

## REFERENCE

Brassicaceae; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Eukaryota; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

## AUTHORS

1 (bases 1 to 1044)  
Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Unterbach, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490

## JOURNAL

Other GSSs: BONK86TR  
Contact: Chris Town

## COMMENT

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TP  
Class: sheared ends.

## FEATURES

Location/Qualifiers  
1..1044

## ORIGIN

/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="T01000DH3"  
/db\_xref="taxon:3712"  
/clone="BONK86"  
/clone\_1b="BO\_1.6.2\_KB\_tot"  
/note="Vector: pHD1; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pHD1 using BstXI linkers"

Query Match 42.1%; Score 614; DB 9; Length 1044;  
Best Local Similarity 78.3%; Pred. No. 6.6e-176;  
Matches 774; Conservative 0; Mismatches 210; Indels 4; Gaps 3;

Qy 444 ACCATGACGGTATGACAGAAAGACCGGCTTAATACGTCAGAGACCGCGGTAT 503  
Db 1004 AGCATGACGGTATCTGGGAAATAGATCGGCTAATCTGTGACAGACCGCGGTAT 945

Qy 504 ACCTAGGATGCAAGCGTTATGGAATTACTGGGCGTTAAAGGTCGACGCGCTTGT 563  
Db 944 ACAGAGAAATGCAAGCGTTATCGGAATGATTGGGCGTTAAAGCTCTGTAGTGGCTTTT 885  
Qy 564 AAGTCAGATGTAATATCCCGGCTTAACTCGGAAATTCGTTGAATCAAAAGCTAG 673  
Db 884 AAGTCCCGCTCAAAATCCAGGGCTCAACCTTGGACAGGCGGTGAAATCAACCACTTG 825  
Qy 624 AGTGGCAGAGGAGGTGAATTCATGTATGACGTGAAATGCTTAAGATATGGAAG 683  
Db 824 AGTACGATGAGGCGAGAGGAAATTCGCGTGAACGCTGAAATGCTAAGATCGAAG 765  
Qy 684 AATATGATGGGCGAAGGAGCGCTCGGGTTAACATGACGCTCATGCAAGAAAGCTG 743  
Db 764 AACACCAAGCGCGAAGACACTTGTGTGGCCGACACTGACACTGAGAGAGAAAGCTAG 705  
Qy 744 GGAGCAAAACAGATTAGATACCTGTGTAGTCAAGCCCTTAAACGATGTCAATGTTGT 803  
Db 704 GGAGCGAATGGATTTAGATACCCAGTATGCTTAAAGCGATGATAGGCGCT 645  
Qy 804 GGGCTTATTAAGCTTG--GTAAAGAACTAACGCGTGAAGTTGACCGCTGGGAGTAC 861  
Db 644 GTGCGTATCGACCGGTGAGTCTGTAGCTTAACGCGTGAATGCCGCTGGGAGTAC 585  
Qy 862 GGTGCGAAGTTAAATCTCAAGAAATTAAGCGGAGACCGGCAAGGCGGTGATTAATGT 921  
Db 584 GTTTCGCAAGATAAATCTCAAGAAATTAAGCGGAGACCGGCAAGGCGGTGATTAATGT 525  
Qy 922 GATTAATTCGATCAACGCGAATAAATCTTAATCTTAATCTTAATCTTAATCTTAATCT 981  
Db 524 GTTAAATTCGATCAACGCGAATAAATCTTAATCTTAATCTTAATCTTAATCTTAATCT 465  
Qy 982 AGATAGATTAGTG-CTTGGGAAACGCTAACAGAGTGTGCAATGCTGTGCACTGCT 1040  
Db 464 AAAGAGAGGAGTGTCTTGGGAAACGCGAACAGAGTGTGCAATGCTGTGCACTGCT 405  
Qy 1041 GTGCGAAGTGTGGTGTAGTCCCGAACAGCGCAACCTTGTCAATTAATGCCATCA 1100  
Db 404 GCGTAAAGTGTGGTGTAGTCCCGAACAGCGCAACCTTGTCAATTAATGCCATCA 345  
Qy 1101 TTGTTGGGCACTTAATGAGACTGCGGTGCAAAACCGAGAGAGTGGGATGACGT 1160  
Db 344 TTGAGTTTGAACCTTAACAGACTGCGGTGATTAACCGAGAGAGTGGGATGACGT 285  
Qy 1161 CAAGTCTCATGAGCCCTTATGAGGTAGGCTTACACAGTAATACATGCGCGTACAGAG 1220  
Db 284 CAAGTCTCATGAGCCCTTATGAGGTAGGCTTACACAGTAATACATGCGCGGACAAAG 225  
Qy 1221 GTTGCCAAACCCGAGGAGAGCTAATCTCAGAAAGCGCGTGTCCGATCGGAGTC 1280  
Db 224 GTGCGAATCCCGCAGGAGTGAAGCTCAAAAAACCGTCTCAAGTTCCGATTGACAGC 165  
Qy 1281 TGCATCTCACTCCGTAAATCGGAAATCGCTAATACGCGGATCAG-CATGTGCGGCTG 1339  
Db 164 TGCATCTCCGCTCATTAAGCCGAAATCGCTAATACGCGGATCAGCGGCTG 105  
Qy 1340 AATACGTTCCCGGCTCTTGTACACACCGCCGTACACCAATGGAGTGGTTTCCAGCA 1399  
Db 104 AATTCTGTTCCCGGCTCTTGTACACACCGCCGTACACCAATGGAGTGGGCAATGCCGA 45  
Qy 1400 AGCAGTAGTCTTAACCGTAAGAGAGGCG 1427  
Db 44 AGTCGTACCTTAACCGCAAGAGAGGCG 17

RESULT 13  
BZ494182/c 1051 bp DNA linear GSS 16-DEC-2002  
LOCUS BONHR54TR\_BO\_1.6.2\_KB\_tot Brassica oleracea genomic clone BONHR54,  
DEFINITION genomic survey sequence.  
ACCESSION BZ494182  
VERSION BZ494182.1 GI:27004681

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
COMMENT

GSS.  
Brassica oleracea  
Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 1051)  
Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Uteback, T.R., Wortman, O.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490  
Other\_GSS8: BONHR54TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.  
Location/Qualifiers  
1..1051  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone\_lib="BONHR54"  
/note="Vector: pHO51; Site 1: BstXI, 1.6-2 kb sheared  
total DNA inserted into pHO51 using BstXI linkers"

ORIGIN

Query Match 42.1%; Score 614; DB 9; Length 1051;  
Best Local Similarity 78.3%; Pred. No. 6.6e-178;  
Matches 774; Conservative 0; Mismatches 210; Indels 4; Gaps 3;

444 ACCCATGACGGTATCGACAGAGAGAGACCGGCTAACTACGTCCAGACCGCGGTAA 503  
1004 AGCAATGACGGTATCTGGGAAATTAAGCATCGGCTAAGTGTGCGACGACCGCGGTAA 945  
504 AGCTAGGCTGCAACCGTTAATCGGAATTAATCTGGCGGTAAAGGGTGGCGCGCGCTTGT 563  
944 ACAGAGAGATGCAAGCGTTATCCGGAATGATTGGCGCGGTAAAGCGTCTGTAGTGGCTTTT 885  
564 AACTCAGATGTAATCCCGGGCTTAACTGGGAAATTTGGTTGAACCTACCAAGCTAG 623  
884 AAGTCGCGCGTCAAAATCCAGGGGCTCAACCTGGACAGCGCGTGGAACTACCAAGCTTG 825  
624 AGTGTGCAGAGGAGGTGGAATTCATGTGTAGCAAGTGAATGCGTAGAGATATGGAAG 683  
824 AGTACGTTAGGGGACAGAGGGAATTTCCGGTGGAGCGGTGAATGCGTAGAGATCGGAAAG 765  
684 AACATGATGGCGAAGGCGAGCTCTCGGTTAACTGACCTGATGACAGAAAGCGTGG 743  
764 AACACCAACGCGCAAAAGCACTGTGCGGCGGACACTGACATGAGAGACGAAAGCTAGG 705  
744 GAGGCAAAACGAGTTGATACCTCGTGAATCCAGGCGCTTAAAGATGTCAAGTTGTT 803  
704 GAGGCGAAATGGGATTAAGATACCCCAAGTATCTTAGCCGTTAAAGATGATATCAGGCGCT 645  
804 GAGGCTTATTAAGGCTTG--GTAAAGAACTAAGCGGTGAAGTTGACCGCTGGGAGTAC 861  
644 GTGCGTATGCAACCGTGCAGTGTCTAGCTAAGCGGTAAAGTATCCGCGCTGGGAGTAC 585  
862 GGTGCGAAGTTAAACTCAAGAGAAATGACGGGAGACCGGACAGCGGTGATTAATGTC 921  
584 GTTCGCAAGAAATAAACTCAAGAGAAATGACGGGAGGCGGACACAGAGGCGTGAAGCATGTC 525  
922 GATTATTCGATGCAACGCGAAACCTTACCTTACCTTGAATGATGAGCAATTTTCTAG 981

Db 524 GTTTAATTCGATGCAAGAGCAAGAACTTACCAAGGAGCTTACATGCGCGGAATCCTCTTG 465  
Qy 982 AGATGATTATGTC-CTTGGGAAACGTTAACACAGAGTCTGATATGCTGTCCAGCTGCT 1040  
Db 464 AAGAGAGGGGAGCTCTTGGGAAACCGGACACAGAGTGTGATGCTGTCTGACCTGCT 405  
Qy 1041 GTGCGAATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGTCAATTAATGGCATCA 1100  
Db 404 GCCGTAAAGTGTGGGTTAAGTCCCGCAACGAGCGCAACCTTGTCAATTAATGGCATCCG 345  
Qy 1101 TTTGTTGGGCACTTAAATGAGACTGCGGTGACAAACCGAGAGAGTGGGATGACGT 1160  
Db 344 TTGAGTTTGAACCTGAAACAGACTGCGGTGATTAAGCCGAGGAAAGGATGAGATGACGT 285  
Qy 1161 CAAGTCTCATGAGCCCTTATGAGGTGAGGCTTACACCTGATATCAATGAGCGCTACAGAG 1220  
Db 284 CAAGTATCATATCCCTTATGCTTGGGCGACACAGTGTCAATTAAGTGGCGGAGCAAAAG 225  
Qy 1221 GTTGCACACCGGAGGGGAGACTAATCTCAGAAAGCGGTGTATGTCGGATCGGAGTTC 1280  
Db 224 GTGCGATCCGCGAGGGTGAAGCTTAACTTCACAAACCCGCTCTCACTTTCGATTGCAAGC 165  
Qy 1281 TGCACCTGACTCCGTGGAATGCGAATGCTAATGCGCGATGAC-CATGTCCGCGTG 1339  
Db 164 TGCACCTGCTGTCATGAAGCCGGAATGCTAATGCGCGATGACCATACGCGCGGTG 105  
Qy 1340 AATPAGTTCCCGGCTTGTGTACACACCGCGCTCACACATGGAGTGGGTTTACACAGA 1399  
Db 104 AATTCGTTCCCGGCTTGTGTACACACCGCGCTCACACTATGGAGTGGGCGCATGCCAA 45  
Qy 1400 AGCAGTAGTCTAACCGTAAAGAGGAGCG 1427  
Db 44 AGTCGTACCTTAACCGCAAGAGGAGCG 17

RESULT 14  
CK280527 889 bp mRNA linear EST 02-AUG-2004  
EST743249 Nicotiana benthamiana mixed tissue cDNA library,  
normalized, full-length Nicotiana benthamiana cDNA clone NBMA117 5'  
end, mRNA sequence.  
CK280527  
CK280527.1 GI:39850196  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

EST.  
Nicotiana benthamiana  
Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamiales; Solanales; Solanaceae; Nicotiana.  
1 (bases 1 to 889)  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
Starkawicz, B., Jin, H. and Baker, B.  
Generation of EST sequences from Nicotiana benthamiana  
Unpublished (2003)  
Other\_ESTs: EST743250  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.  
Location/Qualifiers  
1..889  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NBMA117"  
/issue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"

FEATURES  
source

/note="Vector: PCWVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from *Nicotiana benthamiana* tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (*Pseudomonas syringae* pv tomato 12 hr; *Xanthomonas campestris* pv *campestris* 12 hr, 18hr; *Pseudomonas syringae* pv *phaseolicola* 18hr, and *Xanthomonas campestris* pv *vesicatoria* 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

## ORIGIN

```

Query Match      41.2%; Score 599.8; DB 7; Length 889;
Best Local Similarity 81.3%; Pred. No. 1.5e-173;
Matches 708; Conservative 0; Mismatches 162; Indels 1; Gaps 1;

QY 1 ATTGAACGCTGGGGGCGGCTTTACATGACATGCAAGTGAAGGCGGACGAGATGCTTGCAT 60
DB 18 ATTGAACGCTGGGGGCGGCTTTACATGACATGCAAGTGAAGGCGGACGAGATGCTTGCAT 77
QY 61 CTGTGGCGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 78 CTGTGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 137
QY 121 ACCGATGAAAGATGTCTTAATACCGATATCTTAAGAGAGAAAGAGAGAGAGAGAG 180
DB 138 ACCGTCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 197
QY 181 GACCTTGGCTTTTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTAC 240
DB 198 GGCCTTGGCTTTTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTAC 257
QY 241 CAAGCGCAGCATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
DB 258 CAAGCGCAGCATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 317
QY 301 GCCGAGATCTCTTACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 318 GTCCAGATCTCTTACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 377
QY 361 CAGCAATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 378 CAGCAATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 437
QY 421 AAAGGTACGGTAATATCGTGAACCATGACGATGACAGAGAGAGAGAGAGAGAGAG 480
DB 438 AGGCAATGACGTAATATCGTGAACCATGACGATGACAGAGAGAGAGAGAGAGAGAG 497
QY 481 TACGTGCGCAGAGCGGCGGTAATACGTAAGGTGCAAGCGTTAATCGAATTTACTGGGCGT 540
DB 498 TCGTGTGCGCAGAGCGGCGGTAATACGTAAGGTGCAAGCGTTAATCGAATTTACTGGGCGT 557
QY 541 AAAGGTGCGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 558 AAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 617
QY 601 TGGCTTTGAAACTCAAAAGCTAGAGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 618 TGCATCCAAAGCTGGCAAGCTAGAGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 677
QY 661 TGAATGCGTAGAGATATGGAAGAACTCGATGGCGAAGGCGGCTCTGGGTTAAGACT 720
DB 678 TGAATGCGTAGAGATATGGAAGAACTCGATGGCGAAGGCGGCTCTGGGTTAAGACT 737
QY 721 GACGCTCATGCAAGAAAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 738 GACACTGAGGTGCAAAAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 797
QY 781 CTAACAGATGCACTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 839
DB 798 GTAAACATGTCAACTAGCCGTGGAGAGCTTGAAGCTTTAGTGGCGGAGCTTAAGCAT 857

```

QY 840 AAGTGAACCGCTGGGAGTACGTCGCAAG 870  
DB 858 AAGTGAACCGCTGGGAGTACGTCGCAAG 888

RESULT 15  
BH647750/c 1054 bp DNA linear GSS 19-FEB-2002  
LOCUS  
DEFINITION  
BOMND31R BO\_2\_3\_KB Brassica oleracea genomic clone BOMND31,  
genomic survey sequence.

ACCESSION  
BH647750  
VERSION  
BH647750.1 GI:18705298  
KEYWORDS  
GSS.

SOURCE  
ORGANISM  
Brassica oleracea

REFERENCE  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

Ayelle M., Haas B.J., Kumar N., Mu H., Xiao Y., Van Aken S.,  
Utterback T.R., Wortman J.R., White O.R. and Town C.D.  
Whole genome shotgun sequencing of *Brassica oleracea* and its  
application to gene discovery and annotation in *Arabidopsis*  
Genome Res. 15 (4), 487-495 (2005)  
15805490

COMMENT  
Contact: Chris Town

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..1054

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="TO1000DH3"

/db\_xref="taxon:3712"

/clone="BOMND31"

/clone\_1lb="BO\_2\_3\_KB"

/note="Vector: pHD1; Site 1: BstXI; 2-3 kb sheared  
genomic DNA inserted into pHD1 using BstXI linkers"

## ORIGIN

```

Query Match      40.9%; Score 596.2; DB 9; Length 1054;
Best Local Similarity 77.6%; Pred. No. 2.1e-172;
Matches 772; Conservative 0; Mismatches 218; Indels 5; Gaps 4;

QY 467 AAGCACCAGCTTAATACGTCGACAGCGCGGTAATACGTAAGGTGCAAGCGTTATCG 526
DB 1054 AAGCACCAGCTTAATACGTCGACAGCGCGGTAATACGTAAGGTGCAAGCGTTATCG 995
QY 527 GAATTAATGCGGCGTAAAGGCTGCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 586
DB 994 GAATTAATGCGGCGTAAAGGCTGCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 936
QY 587 CTTAACCTGGGAATGCGTTGAAACTCAAAAGCTAGAGTGTGCGCAGAGGAGGTGAAT 646
DB 935 CTTAACCTGGGAATGCGTTGAAACTCAAAAGCTAGAGTGTGCGCAGAGGAGGTGAAT 876
QY 647 TCCATGTGACAGTGAATGCGTGAAGATATGGAAGAACTCGATGGCGAAGGCGGAGCT 706
DB 875 TTCGGTGGAGCGGTGAATGCGTGAAGATATGGAAGAACTCGATGGCGAAGGCGGAGCT 816
QY 707 CTTGGGTTAATCACTGACGCTCAAGCAAGAAAGGCTGGGAGCAAGAGATTGATGCC 766
DB 815 GCTGGGCGGACCTGACCTGAAGAGAGAAAGCTAGGAGGAGCAAGAGATTGATGCC 756
QY 767 TGTAGTCCAGCGCCTTAAGAGATGCACTAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 824
DB 755 CAGTAGTCTTAGCCGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 696

```

QY 825 CGAAGCTAACGGCTGAAAGTTACCGCCTGGGGAGTACGGTCCGAAAGTTAAACTGAAG 884  
DB 695 TGTAGCTTAACGGGTTAAAGTATCCCGCTGGGGGTACGTTCCGAAAGTAAAGAACTCAAG 636  
QY 885 GAATTGACGGGGGACCGGCAACAGCGGTGATTAAGATTAATTGATGCAACGGGAA 944  
DB 635 GAATTGACGGGGGACCGGCAACAGCGGTGATTAAGATTAATTGATGCAACGGGAA 576  
QY 945 AACCTTACCTACCTCTTGAATGATGAGAAATTTCTAGAGATGATTAAGTG-CTTCGGAA 1003  
DB 575 AACCTTACCAAGGCTTGAATGATGCGGAAATCTTGAAGAGAGGGGTGCTTCGGAA 516  
QY 1004 CCGTAAACAGAGTGTGATGATGCTGTGATGATGCTGTGATGATGATGATGATGATG 1063  
DB 515 CCGGACACAGAGTGTGATGATGCTGTGATGATGATGATGATGATGATGATGATGATG 456  
QY 1064 CCGCAACAGAGCGGACCTTGTGATTAATGATTAATGATTAATGATTAATGATTAATG 1123  
DB 455 CCGCAACAGAGCGGACCTTGTGATTAATGATTAATGATTAATGATTAATGATTAATG 396  
QY 1124 CTGCGCGTGAACAACCGGAGAGAGTGGGATGACGTCAAGTCCATGAGCCCTTATGG 1183  
DB 395 CTGCGCGTGAACAACCGGAGAGAGTGGGATGACGTCAAGTCCATGAGCCCTTATGG 336  
QY 1184 TAGGGCTTCAACAGTAAATCAATGGGCGCTGACAGAGGTTGCCAACCCCGAGGGGAG 1243  
DB 335 CTGGGCGACACAGTCTACATGATGGCGGACAAAGGTCGCGATCCCGAGGAGTGAAC 276  
QY 1244 TATCTCAGAAAGCGGCTGTGATGCTGGATCGGATCGGATCGGATCGGATCGGATCG 1303  
DB 275 TATCTCAGAAAGCGGCTGTGATGCTGGATCGGATCGGATCGGATCGGATCGGATCG 216  
QY 1304 GAATCGCTGATTAATCGCGATCAG-CATGTCGGCGTGAATACGTTCCGGGCTTTGACA 1362  
DB 215 GAATCGCTGATTAATCGCGATCAGCATACGGGCGTGAATGCTTCGCGGCTTTGACA 156  
QY 1363 CACCGCGCTGACACATGAGAGTGGTTTACCAAGACAGTACTTAACTGTAAGGA 1422  
DB 155 CACCGCGCTGACACATGAGAGTGGTGGACCTGACATGCGCAAGTCCTTAAACGCAAG 96  
QY 1423 GGGCGCTTGCACGAGTGAATTCATGACTGGGGTG 1457  
DB 95 GGGCGCTTGCACGAGTGAATTCATGACTGAGTG 61

RESULT 16  
CK280528 884 bp mRNA linear EST 02-AUG-2004  
LOCUS EST743250 Nicotiana benthamiana mixed tissue cDNA library,  
DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NBMA117 5',  
end, mRNA sequence.  
ACCESSION CK280528  
VERSION CK280528.1 GI:39850200  
KEYWORDS EST.  
SOURCE Nicotiana benthamiana  
ORGANISM Nicotiana benthamiana  
Bukayocsa; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamids; Solanales; Solanaceae; Nicotiana.  
1 (bases 1 to 884)  
Buell, C.R., Hart, A., Zemann, V., Karamycheva, S.A., Day, B.,  
Staskawicz, B., Jin, H. and Baker, B.  
Generation of EST sequences from Nicotiana benthamiana  
Unpublished (2003)  
Other ESTs: EST743249  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES  
source  
Location/Qualifiers  
1..884  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NBMA117"  
/tissue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-TonA"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"  
/note="Vector: PCWVSp06.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Nicotiana benthamiana  
tissues that include callus, roots from liquid culture  
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
cold-stressed leaves (5 C, 3 hr, 6hr), and pathogen  
challenged leaves (Pseudomonas syringae pv tomato 12 hr;  
Xanthomonas campestris pv campestris 12 hr, 18hr;  
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas  
campestris pv vesicatoria 18hr). RNA was isolated from  
these tissues and pooled in approximately equal molar  
amounts."

ORIGIN  
Query Match 40.9%; Score 595.8; DB 7; Length 884;  
Best Local Similarity 81.2%; Pred. No. 2,7e-172;  
Matches 704; Conservative 0; Mismatches 162; Indels 1; Gaps 1;  
QY 1 ATTTGAACGCTGGCGCATGCTTTTACATGCAATGCAACGGCAGCAGATGCTTGAT 60  
DB 18 ATTTGAACGCTGGCGCATGCTTTTACATGCAATGCAACGGCAGCAGATGCTTGATC 77  
QY 61 CTGGTGGGAGTGGGGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATG 120  
DB 78 CTGGTGGGAGTGGGGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATG 137  
QY 121 ACGCATCGAAAGATGTGCTTAATACCGCATATCTCTTAAGAGGAAAGCAGGGGATCGAA 180  
DB 138 ACGCTCGAAAGAGAGCTTAATACCGCATATCTCTTAAGAGGAAAGCAGGGGATCG 197  
QY 181 GACCTTGGCGCTTTTGAAGCGCGATGCTGATTAAGTAACTTGAATGATGATGATG 240  
DB 198 GGCCTTGGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 257  
QY 241 CAAAGCGAGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
DB 258 CAAAGCGAGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 317  
QY 301 GCCCAAGCTCTTACGAGAGAGCAGAGTGGGGAATTTTGAACAATGGGCGCAAGCTGATC 360  
DB 318 GTCCAGATCTCTTACGAGAGAGCAGAGTGGGGAATTTTGAACAATGGGCGCAAGCTG 377  
QY 361 CAGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
DB 378 CAGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 437  
QY 421 AAGGTTACGATTAATTAATCGTGAACCAATGACAGTATGACAGAGAGCAACGGCTAAC 480  
DB 438 AAGGTTACGATTAATTAATCGTGAATTTTGAAGTTACGACAGAAATGACACGGCTAAC 497  
QY 481 TAGGTGCCAGACCGCGGTAAATAGTGAAGGTGCAAGCTTAAATGCAATTAATGAGCGCT 540  
DB 498 TCTGTGCCAGACCGCGGTAAATAGTGAAGGTGCAAGCTTAAATGCAATTAATGAGCGCT 557  
QY 541 AAGGTTGCCAGACCGCGGTAAATAGTGAAGGTGCAAGCTTAAATGCAATTAATGAGCG 600  
DB 558 AAGGTTGCCAGACCGCGGTAAATAGTGAAGGTGCAAGCTTAAATGCAATTAATGAGCG 617  
QY 601 TCGCTTTGAATTAATCAAAAGCTGATGATGATGATGATGATGATGATGATGATGATG 660  
DB 618 TGCATTCCAAAATTTGCAAACTGATGATGATGATGATGATGATGATGATGATGATG 677  
QY 661 TGAATATGCTAGATATGAGAGAAATGATGATGATGATGATGATGATGATGATGATG 720

Db 678 TGAATGCTGATGATATGGAAGAACCAAGTGGCGAAGGCGACCTGAGCTGATACT 737  
Qy 721 GAGCTCATGACAGAAAGCGTGGGAGCAACAGATTTAGTACCTGTAGTCAAGCC 780  
Db 738 GACACTAGAGTGGCAAGGCGTGGGAGCAACAGATTTAGTACCTGTAGTCAAGCC 797  
Qy 781 CTAAACGATGTCAACTAGTTGTTGGGCTTATTAG-GCTTGTAAAGCAAGCTAACGCGTG 839  
Db 798 GTAAACGATGTCAACTAGCTTGGAGCCTTGAAGCTTATAGTGGCGCAGCTAACGAT 857  
Qy 840 AAGTTGACCGGCTTGGGAGTACGCGTG 866  
Db 858 AAGTTGACCGGCTTGGGAGTACGCGCG 884

RESULT 17  
BZ685785/c 899 bp DNA linear GSS 05-FEB-2003  
LOCUS PUBD154TD ZM\_0.6\_1.0\_KB Zea mays genomic clone ZMMBtra032112,  
DEFINITION genomic survey sequence.  
ACCESSION BZ685785  
VERSION BZ685785.1 GI:28245311  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 899)  
AUTHORS WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,  
Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennerzen,J.  
TITLE Maize Genomics Consortium  
JOURNAL Unpublished (2003)  
COMMENT Contact: Cathy WhiteJaw  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteJaw@tigr.org  
Seq primer: TP  
Class: sheared ends.  
FEATURES  
source Location/Qualifiers  
1..899  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBtra032112"  
/clone\_id="ZM\_0.6\_1.0\_KB"  
/note="Vector: pCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
cot selected genomic DNA library"

ORIGIN  
Query Match 40.6%; Score 592; DB 9; Length 899;  
Best Local Similarity 80.3%; Pred. NO. 4e-171;  
Matches 742; Conservative 0; Mismatches 155; Indels 27; Gaps 3;

Qy 318 AGGCAAGAGTGGGAAATTTGGACAATGGGCGCAAGCCTGATCCGACATGCCGCGTGA 377  
Db 899 AGGCAAGAGTGGGAAATTTGGACAATGGGCGCAAGCCTGATCCGACATGCCGCGTGA 840  
Qy 378 TGAAGAAGGCTTGGGTTGTAAAGCTCTTTCAGTCCGAGAGAAAGGTTACGTAATA 437  
Db 839 TGATTAAGGCTTGAAGGTTGTAAAGCTCTTTCAGTCCGAGAGCAATA----- 795  
Qy 438 ATCGTGAACCATGATCGATCGACAGAAAGACACCGGCTTAAGTCTGCGCAGCAGCCGC 497  
Db 794 -----ATGACGATACCGGAAGATTAAGCCCGGCTTAAGTCTGCGCAGCAGCCGC 745  
Qy 498 GGTAAATACGTAGGTGCAAGCGCTTAATCGGAATTAAGTGGGCTTAAGGCTGCGCAGCGG 557

Db 744 GGTAAATACGAAGGGGGCTAGCGTTGCTCGGAATCACTGGGCTTAAGGGCGCGTAGCGG 685  
Qy 558 CTTTGAATGTCAGATGTGAATATCCCGGCTTAAGCTGGGAATTTGGCTTGAATCAACA 617  
Db 684 CCATTCAAGTCCGGGGGTAAGAGCTGTGCTCAACACAAATTTGCTTGAATATGTTT 625  
Qy 618 AGCTAAGAGTGTGCAAGGAGGTTGAATTCATGTGTAGCAGTGAATATCGTAGAGATA 677  
Db 624 GGCTTGAATTTGTAGAGTTGAGTTGGTGTGAATCTGCAGATGTAGAGTGAATTCGTAGAT 565  
Qy 678 TGAAGAAGCATTCAGTGGCGAAGGCAAGCCTCTGGGTTAACTGACGCTCATGACGA 737  
Db 564 CGCAAGAACCAAGTGGCGAAGGCGGCAACTGACCAACACTGACGCTGAGGCGCAAA 505  
Qy 738 GCGTGGGAGCAACAGAGTTAATACCTGTGTATGTTTCAAGCCCTTAACGATGCACTA 797  
Db 504 GCGTGGGAGCAACAGAGTTAATACCTGTGTATGTTTCAAGCCCTTAACGATGATCTA 445  
Qy 798 GTTGTGGGCTTATTAAGCTGTGTAAAGAGCTAACGCGTGAAGTTGACCGCTGGGGA 857  
Db 444 GCTGTGGGCTTATTAAGCTGTGTAAAGAGCTAACGCGTGAAGTTGACCGCTGGGGA 385  
Qy 858 GTACGCTGCGAAGTAAATCTCAAGAGATTGAACGCGGAGACCCGACAGCGGTGATTA 917  
Db 384 GTACGCTGCGAAGTAAATCTCAAGAGATTGAACGCGGAGACCCGACAGCGGTGATTA 325  
Qy 918 TGTGATTAATTTGATGCAACGCGAAGAACTTAACCTTGAACATGTATGCAATTTT 977  
Db 324 TGTGATTAATTTGATGCAACGCGAAGAACTTAACCTTGAACATGTATGCAATTTT 265  
Qy 978 CTAGAGAT-AGATTAGTCTTGGGAGCGTAAACAGAGTGTGATGCTGCTGCTGAGC 1036  
Db 264 GGAAGAGTCCGGGGTCTTCCCTTGGGAGCGGACAGACAGTGTCTGCTGCTGAGC 205  
Qy 1037 TCGTGTGATGAGATGTTGGGTTAAGTCCCGCAACGAGCGAACCCTTGTCTTAATTGCC 1096  
Db 204 TCGTGTGATGAGATGTTGGGTTAAGTCCCGCAACGAGCGAACCCTTGTCTTATGCC 145  
Qy 1097 ATCAATTTGTTGGGCACTTAATGACACTGCGGTTGCAAAACG-GAGGAAGTGGGAT 1155  
Db 144 ATCAATTTGTTGGGCACTTAATGACACTGCGGTTGTAAGCCCGAGGAAGTGTGAT 85  
Qy 1156 GACGTCAAGTCCATAGGCGCTTAATGGGTTGAGGCTTCAACGTAAATCAATGGCGGTAC 1215  
Db 84 GACGTCAAGTCCATAGGCGCTTAATGGGTTGAGGCTTCAACGTAAATGGCGGTAC 25  
Qy 1216 AGAGGTTGCGCAACCGCGAGGGG 1239  
Db 24 AGTGGAGCGCAACCGCGAGGTG 1

RESULT 18  
BZ440868 974 bp DNA linear GSS 13-DEC-2002  
LOCUS BONKTS1TR BO\_1.6\_2\_KB\_tot Brassica oleracea genomic clone BONKTS1,  
DEFINITION genomic survey sequence.  
ACCESSION BZ440868  
VERSION BZ440868.1 GI:26698161  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; euroside II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 974)  
AUTHORS Aylee,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Uteerback,T.R., Wortman,J.R., White,O.R. and Town,C.D.  
TITLE Whole genome shotgun sequencing of Brassica oleracea and its  
JOURNAL application to gene discovery and annotation in Arabidopsis  
PUBMED Genome Res. 15 (4), 487-495 (2005)  
15805490  
COMMENT Other GSSs: BONKTS1TF  
Contact: Chris Town



[illegible]

FEATURES		Seq primer: TR	
SOURCE		Class: sheared ends.	
Location/Qualifiers			
1..951			
/organism="Brassica oleracea"			
/mol_type="genomic DNA"			
/strain="TOL000DH3"			
/db_xref="taxon:3712"			
/clone="BOMFE30"			
/clone_1db="BO_2_3_KB"			
/note="Vector: pHD1; Site 1: BclXI; 2-3 kb sheared genomic DNA inserted into pHD1 using BclXI linkers"			
ORIGIN			
Query Match	39.4%	Score 574;	DB 9; Length 951;
Best Local Similarity	77.8%	Pred. No. 1.5e-165;	
Matches 731;	Conservative 0;	Mismatches 205;	Indels 4; Gaps 3;
Db	502	ATACGTAGGGGCGAAGCGTTAATCGGAATTAAGTGGGCGTAAAGGGTGCAGAGCGGCGCTT	561
6	ATACAGAGGATCGACAGCGTTATCCGGAATGATTTGGCGTTAAAGGCTCTGTAAGTGGCTTT	65	
Qy	562	GTAAGTCAGATGTGAAATCCCGGGCTTAACTTGGGAATTGCGTTGAAACTACAAAGCT	621
Db	66	TTTAAGTCGGCGTTAAATCCAGGGCTCAACCTCGACAGCGGGTGAATCAACCAAGCT	125
Qy	622	AGAGTGTGGCAGAGGAGGTGGAATTCATGTGTGACAGTGAATTGGGTAGAGATATGCA	681
Db	126	TGAGTACGGTAGGGGCGAGAGGAATTTCCGGTGGAGCGGTGAATAATGGGTAGAGATCGGAA	185
Qy	682	AGAACATCCGATGGGGAAGGACGCTCCGGGTTTAACTGACGCTCATGACAGAAAGCGT	741
Db	186	AGAACACCAACGCGGAAAGCACTTGCTGGGCGCACCTGACCTGAGAGACGAAAGCTA	245
Qy	742	GGGAGAGCAACAGGATTAGATATCCCTGTAAGTCCAGCCCTTAAACGATGTCAACTAGTTG	801
Db	246	CGGAGAGGAATGGGATTAGATATCCCGAGTATCTTACCGCTTAAACGATGTACTAGGCG	305
Qy	802	TTGGGCGCTTATTAAGGCTTG--GTACGAACTTAAACGGGTGAAGTTGACCCGCTGGGAGT	859
Db	306	CTGTGCGTATCGACCCCGTCAGTCTGTAGCTTAAACGGGTTAAGTATCCCGCTGGGAGT	365
Qy	860	ACGGTCCCAAGATTAAATCTCAAGGAATTGACGGGGGACCGGCAACAGCGGTGATTAATG	919
Db	366	ACGTTCCGAAGAAATGAATCTAAAGAAATGACGGGGGCCCGCACAGCGGTGAGCATG	425
Qy	920	TGATTAATTGATGACAAACGGAATAAATCTTACCTACCTTGAAGATGAGCAATTTTCT	979
Db	426	TGGTTAATTGATGACAAACGGAATAAATCTTACCAAGGCTTGAATGATCGCGGAATCTCT	485
Qy	980	AGAGATAGATTAGTG-CTTGGGAAACGCTTAAACAGAGTGTGCAATGGCTGTCAAGTTC	1033
Db	486	TGAAGAAGAGGGGTGCTTCGGGAAACGCGGACACAGGGTGTGCAATGGCTGTCAAGTTC	545
Qy	1039	GTGTGCGAAGATGTTGGGTTAAGTCCCGCAACGAGCGGAACCCCTTGTCATTAATTGGCAT	1098
Db	546	GTGCGCTTAAGGTGTGGGTTAAGTCCCGCAACGAGCGGAACCCCTGTGTTAAGTTGCCAC	605
Qy	1099	CATTGGTGGGCACTTTAATGAGACGTGCGGTGCAAAACGAGAGGAAGTGGGGATGAC	1158
Db	606	CGTTAGTTTGAACCTCTGAACAGATCTGCGGTGTAAGCCGAGAGGAAGTGAAGATGAC	665
Qy	1159	GTCAAGTCTCATGGCCCTTATGGGTAGAGGCTTCAACAGTAATCAATGGCGCTGACGA	1218
Db	666	GTCAAGTATCATGCCCTTATGCTGCGGAGCAACAGGTCTCAATAGGCCCGGAGCAAA	725
Qy	1219	GGGTTGCCAATCCCGAGAGGAGGAGCTAATCTCAAGAAAGCGGTGTGATCCGGAGTCGAG	1278
Db	726	GGGTGCGCATCCCGAGAGGAGGAGCTAATCTCAAGAAACCCGTCTCAATTTGGAATTCGAG	785
Qy	1279	TCTGCAATCGATCCCGTGAAGTGGGAATGCTGTAGTAATCGCGGATCAG-CATGTCGCG	1337
Db	786	GCTGCAATCTCCCTGCAAGAGCCGGAATGCTGTAGTAATCCCGGTACGCATACGCGG	845



QY 1338 TGAATACGTTCCGGGCTTTGTACACACGCGCCGTACACCATGGAGTGTTTACCA 1397  
|||||  
DB 846 TGATTCGTTCCCGGGCTTGTACACACCGCCGTACACTATGGAGCTGCCATGCC 905  
|||||  
QY 1398 GAAGCAGTAGTCTTAACCGTTAGAGAGCGCTTGCCACG 1437  
|||||  
DB 906 GAAGTCGTTACCTTAACCGCAAGAGGCGGTCCGAGG 945  
|||||

RESULT 21  
CO159272 787 bp mRNA linear EST 18-JUN-2004  
LOCUS FLU1\_12\_C07\_A029 Root flooded Pinus taeda cDNA clone  
DEFINITION FLU1\_12\_C07\_A029 5', mRNA sequence.  
CO159272  
CO159272.1 GI:48929813  
EST.  
SOURCE Pinus taeda (loblolly pine)  
ORGANISM Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
1 (bases 1 to 787)  
Pratt, L., Cordomier-Pratt, M.-M., Lorenz, W.W., Zimmermann, C.,  
Johnson, H., Anfuso, C., Kamran, D., Chhabra, D. and Dean, J.F.D.  
A loblolly pine (Pinus taeda) EST database from flooded roots  
Unpublished (2004)  
Other ESTs: FLU1\_12\_C07\_b1\_A029  
Contact: Cordomier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz (School of  
Forest Resources, University of Georgia); plant material prepared  
by Craig Zimmermann (School of Forest Resources, University of  
Georgia) using rooted cuttings provided by the Forest Biology  
Research Cooperative (FBRC) and the CCLONES project at the  
University of Florida; sequencing done in the Laboratory for  
Genomics and Bioinformatics, University of Georgia. Sequence ends  
have been trimmed to exclude vector and regions below Phred quality  
16. Three-prime sequences are presented as their reverse complement  
and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACGCTATGACC).  
Location/Qualifiers  
1..787  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/db\_xref="taxon:3352"  
/clone="FLU1\_12\_C07\_A029"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_1fb="Root flooded"  
/note="Organ: root; Vector: pSL180; Site 1: EcoRI;  
Site 2: XhoI. The library was prepared from polyA+ RNA  
from the roots of 1-year-old loblolly pine (Pinus taeda)  
cuttings that were rooted and then planted in washed sand.  
Prior to harvesting tissues for RNA isolation, the rooted  
cuttings were maintained for 27 days (April 2003) under  
ambient conditions in a local greenhouse. They were kept  
on a weekly regimen of 0.5x nutrient-complete Hoagland's  
solution and supplemented with additional water sufficient  
to maintain 5% soil moisture content. Pots holding the  
rooted cuttings were fully submerged in water for 24 hours  
prior to harvest of the roots for RNA isolation.  
Double-stranded cDNA was cloned unidirectionally into  
pSL180. Inserts can be excised with EcoRI (5' end) and  
XhoI (3' end)."

ORIGIN  
Query Match 39.3%; Score 572.4; DB 7; Length 787;

Best Local Similarity 85.0%; Pred. No. 4,5e-165;  
Matches 655; Conservative 0; Mismatches 107; Indels 9; Gaps 1;  
QY 4 GAACGTGGCGGCAGCTCTTACACATGCAAGTGCAGACCGGACACAGGATCTTGATCG 63  
|||||  
DB 26 GAACGTGGCGGCAGCTCTTACACATGCAAGTGCAGACCGGACAGT-----TAAG 76  
|||||  
QY 64 GTGGCAGAGTGGGAGCAGCGGTGATATGATCGCAACCTATCCAGAGAGGCGGTACG 123  
|||||  
DB 77 CTGACAGATGGCGAAGCGGTGATATATATGGAACGTCGACGTCGCGGGGATACG 136  
|||||

QY 124 CATCGAAGATGTCTTAATACCGCATATACCTTAAGAGAGAAAGCAGGGGATCGAAAGC 183  
|||||  
DB 137 CAGCGAAGACTGTCTTAATACCGCATATACCTTAAGAGAGAAAGCAGGGGATCGCAAGAC 196  
|||||

QY 184 CTTCGCTTTTGGAGCGCGCGCATGCTGATTTAGCTGATGTTGGGTAAAGGCTTACCA 243  
|||||  
DB 197 CTGCGCGATTTGAGAGCGCGCATATCAGATTAGCTGATGTTGGGTAAAGGCGCACCA 256  
|||||

QY 244 GCGCAGATCAGTATGTTGGTCTGAGAGAGACAGCCACACTGGGACTGAGACAGCGCC 303  
|||||  
DB 257 GCGCAGATCTGTAGCTGATCTGAGAGAGAGACAGCCACACTGGGACTGAGACAGCGTC 316  
|||||

QY 304 CAGACTCTTACCGGAGAGCAGCAGTGGGAAATTTGGACATGCGCGCAAGCTGATCCAG 363  
|||||  
DB 317 CAGACTCTTACCGGAGAGCAGCAGTGGGAAATTTGGACATGCGCGCAAGCTGATCCAG 376  
|||||

QY 364 CAATCGCGGTGAGAGAGAGAGCGCTTGGGTTTAAAGCTCTTTCAGTGCAGAAAGAA 423  
|||||  
DB 377 CAATCGCGGTGAGAGAGAGAGCGCTTGGGTTTAAAGCTCTTTCAGTGCAGAAAGAA 436  
|||||

QY 424 GGTACGTAATTAATCTGACCATGACGATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 483  
|||||  
DB 437 GGTACGCTTAATTAATCTGACCATGACGATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 496  
|||||

QY 484 GTGCCAGACCGCGGTAAATAGTAAAGGTGCAAGCTTATGGAATTAATCTGGCGTTAA 543  
|||||  
DB 497 GTGCCAGACCGCGGTAAATAGTAAAGGTGCAAGCTTATGGAATTAATCTGGCGTTAA 556  
|||||

QY 544 GGTGGCGAGCGCGGTAAATAGTAAAGGTGCAAGCTTATGGAATTAATCTGGCGTTAA 603  
|||||  
DB 557 GGTGGCGAGCGCGGTAAATAGTAAAGGTGCAAGCTTATGGAATTAATCTGGCGTTAA 616  
|||||

QY 604 GTTGAATCTAAGAGCTGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663  
|||||  
DB 617 ATTGAGATCTGATGTCAG 676  
|||||

QY 664 AATGCGTAGATATAG 723  
|||||  
DB 677 AATGCGTAGATATAG 736  
|||||

QY 724 GCTCATGACAGAAAGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774  
|||||  
DB 737 GCTCATGACAGAAAGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 787  
|||||

RESULT 22  
CK290947 851 bp mRNA linear EST 02-AUG-2004  
LOCUS CK290947  
DEFINITION EST153661 Nicotiana benthamiana mixed tissue cDNA library.  
normalised, full-length Nicotiana benthamiana cDNA clone NBMBY50 5'  
end, mRNA sequence.  
CK290947  
CK290947.1 GI:39870911  
EST.  
SOURCE Nicotiana benthamiana  
ORGANISM Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamiales; Solanales; Solanaceae; Nicotiana.  
1 (bases 1 to 851)  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
Stebkiewicz, B., Jin, H. and Baker, B.



TITLE Generation of EST sequences from *Nicotiana benthamiana*  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Robin Bueli  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from the University of Arizona Genomics  
 Institute via <http://genome.arizona.edu/orders/>.  
 Seq primer: ATT TAG GCG ACA CTA TAG.  
 Location/Qualifiers

## FEATURES

source  
 1..851  
 /organism="Nicotiana benthamiana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4100"  
 /clone="NBMBY50"  
 /tissue\_type="abiotic and biotic stress-treated leaves,  
 callus tissue and root tissue"  
 /lab\_host="DH10B-TonA"  
 /clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
 library, normalized, full-length"  
 /note="Vector: PCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 tissues that include callus, roots from liquid culture  
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen  
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;  
 Pseudomonas campestris pv campestris 12 hr, 18hr;  
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas  
 campestris pv vesicatoria 18hr). RNA was isolated from  
 these tissues and pooled in approximately equal molar  
 amounts."

## ORIGIN

Query Match 39.2%; Score 571; DB 7; Length 851;  
 Best Local Similarity 80.9%; Pred. No. 1.2e-164;  
 Matches 690; Conservative 0; Mismatches 160; Indels 3; Gaps 2;  
 4 GAAAGCTGGCGGCTTTTACATCAGTCAAGTCGACGCGACGCGATCTTGATCTG 63  
 1 GAAAGCTGGCGGCGCTTAAACATCAGTCAAGTCGACGCGACGCGATCTTGATCTG 60  
 64 GTGGCGAGTGGCGGCGGAGTAAATGATCGAAGTATCCGAAAGAGGGGGGTAAAG 123  
 61 GTGGCGAGTGGCGGCGGAGTAAATGATCGAAGTATCCGATGCTGGTGGGAGTAAAG 120  
 124 CATCGAAGATGTGCTAATCCGATATCTAAGGAGAAAGCAGGGGATCGAAAGAC 183  
 121 CTGGAAACGAGCGCTAATACGATACGCTCTACGGAGAAAGCAGGGGATCTTGAGGC 180  
 184 CTGGCGCTTTTGGAGCGCGCGATGTGATTAAGTGGTGGGTTAAAGGCTTACCAA 243  
 181 CTGGCGCTTATCAGTGAAGCTTACGATTAAGTGGTGAAGTAAATGGTCAACCA 240  
 244 GGGCAGCATGATGATGTTGCTGAGAGACGACCACTCTGGACTGAGACAGGCGC 303  
 241 GGGCAGCATGATGATGTTGCTGAGAGATGATCACTGGAATGAGACAGGCTC 300  
 304 CAGACTCTTAAAGGAGAGAGAGGGAATTTGGAACAAGGGCGCAAGCTGATCCAG 363  
 301 CAGACTCTTAAAGGAGAGAGAGGGAATTTGGAACAAGGGCGCAAGCTGATCCAG 360  
 364 CAATGCGCGGTGATGAAGAGGCTTGGGTTGTAAGCTCTTCACTGAGAGAA 423  
 361 CCAATGCGCGGTGATGAAGAGGCTTGGGTTGTAAGCTTCACTGAGAGAA 420  
 424 GGTAAAGTAAATATCTGACCCATGACGGTATGACAGAAAGACCGGCTTAATAC 483  
 421 GCAATTAATTAATGATGATGTTTGAAGTAAACGAGAAATAGCAACCGGCTTAATAC 480  
 484 GTGCCAGAGAGCGCGGTAAATAGTATGGTGAAGGCTTAATCGGAATTAATCGGGGTAA 543  
 481 GTGCCAGAGAGCGCGGTAAATAGTATGGTGAAGGCTTAATCGGGGTAA 540

QY 544 GGGTGCAGAGCGGCGCTTGAATGATGTAATCCCGGAGCTTAACTGGGAATTGC 603  
 DB 541 --GCGGTAGGTGGTTTGTAAATGATGTAATCCCGGAGCTTAACTGGGAATTGC 598  
 QY 604 GTTTGAATTAACAAAGCTGAGTGGCAGAGGAGTGGAAATTCATGTATGCACTGA 663  
 DB 599 ATCCAAACTGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 658  
 QY 664 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 723  
 DB 659 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 718  
 QY 724 GCTCATGACGAAAGGCTGGGAGCAACAGATTAATACCTGATGATGATGATGATGATG 783  
 DB 719 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 778  
 QY 784 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 842  
 DB 779 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838  
 QY 843 TTGACCGGCTGGG 855  
 DB 839 TTGACCGGCTGGG 851

RESULT 23  
 BZ459292 977 bp DNA linear GSS 13-DEC-2002  
 LOCUS BONKQ28TF.D0.1.6.2 KB.tot Brassica oleracea genomic clone BONKQ28.  
 DEFINITION genomic survey sequence.  
 ACCESSION BZ459292 GI:26740661  
 VERSION BZ459292.1 GI:26740661  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 977)  
 Ayele,M., Haas,B.D., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
 Utecherback,T.R., Wortman,J.R., White,O.R. and Town,C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 15805490  
 PUBMED  
 JOURNAL Ocher GSSes: BONKQ28TR  
 CONTACT Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TF  
 Class: sheared ends.  
 Location/Qualifiers

## FEATURES

source  
 1..977  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO100DH3"  
 /db\_xref="taxon:3712"  
 /clone="BONKQ28"  
 /clone\_lib="BO\_1.6.2\_KB.tot"  
 /note="Vector: pHO51; Site 1: BstXI, 1.6-2 kb sheared  
 total DNA inserted into pHO51 using BstXI linkers"

## ORIGIN

Query Match 38.7%; Score 563.8; DB 9; Length 977;  
 Best Local Similarity 75.8%; Pred. No. 2.2e-162;  
 Matches 747; Conservative 0; Mismatches 212; Indels 26; Gaps 3;  
 347 GCGAAGCTGATCCAGCAATGCGCGGTGATGAAGAGGCTTGGGTTGTAAGCTCT 406

Db	1	GCMAAAGCTGACGGAGCAATGCCGCGGAGGATATAAGGCTACGGGTCTGTAACTTCT	60
Qy	407	TTCACTGAGAAAGAAAGGTTACGGTAATATATGTGACCATGACCGTATCGACAAAG	466
Db	61	TTTCCAGAGAAAGAG-----CAATACCGTATCTGGGAAAT	97
Qy	467	AAGCACCGGCTTAACATGCTGCGACAGACCGCGGTAAATCTAGAGGTGCAACCGTTAAATCG	526
Db	98	AAGCATCGGCTTAATCTGTGCGACAGACCGCGGTAAATACAGAGATGCAAGCGTTATCCG	157
Qy	527	GAATTACTGGGCGTAAAGGATGCCAGGCGGCTTTGTAAGTCAGATGTAAATCCCGGG	586
Db	158	GAATGATTTGGCGTAAAGCGTCTGATGGTGGCTTTTAAATGTCGCGCGTCAATCCAGGG	217
Qy	587	CTTAACCTGGGAATTCGTTGAAACTTAACAAGTGAAGTGTGGCAGAGGAGGTGAAAT	646
Db	218	CTCAACCTGGAACAGGCGGTGGAAACTCAAGCTTGATCGGTAGGGGAGAGGGAAAT	277
Qy	647	TCCATGCTGACAGTGAATGCGTATGAGATATGGAAGAACATCGATGGCGAAGGACAGCT	706
Db	278	TTCCGGTGGACGGTGAATGCGTATGAGATCGAAGAAACCAACGCGGAAGACATCT	337
Qy	707	CCTGGGTTAACATGACCGCTCATGACGAAAGCGTGGGAGCAACAGATTAATACC	766
Db	338	GCTGGGCGGACATGACATGAGAGCAAAAGCTAGGGGACGAATGGATTAATATACC	397
Qy	767	TGTAATCTCAACGCCCTTAAACGATGCACTATGTTTGGGCTTATTAAGCTTG--GTAA	824
Db	398	CAGTATGTCCTTAGCCGTAAACGATGGATATCTAGGCGCTGTGGTATCGACCGTCAAGTC	457
Qy	825	CGAAGCTTAACGCTGAAGATTGACCGGCTGGGAGATACGGTCGCAAGATTAAACTCAAG	884
Db	458	TGTAGCTAACCGTTAATGATATCCCGCTGGGGAATGACTTGCAGAAATGAATCTCAAG	517
Qy	885	GAATTTGACGGGACCCGCAACAGCGGTGATATGTGATTAATTCGATGCAACGCGAA	944
Db	518	GAATTTGACGGGACCCGCAACAGCGGTGACATATGTGTTAATTCGATGCAAGCGAAG	577
Qy	945	AACCTTAACCTTGAATGATGACGCAATTTTCTAGAGATAGATATGTC--CTTCGGGAA	1003
Db	578	AACCTTAACAGGGCTTGACATGACGCGCAATCTCTTGAAACAGAGGGGTGCTTCGGGAA	637
Qy	1004	CGCTTAACACAGGTCTGTGATGGTGTGCTGACGTGTCGTGAGATTTGGGTTAATGTC	1063
Db	638	CGCGGACACAGTGTGTGATGGTGTGCTGACGTGTCGTGAGATTTGGGTTAATGTC	697
Qy	1064	CCGCAACGAGCGCAACCTTGTCAATTAATGCCATCAATTTGGTGGCACTTTAATGAGA	1123
Db	698	CCGCAACGAGCGCAACCTTGTGTATTAATTTCCACCTGTGATTTGGAAACCTTAACAGA	757
Qy	1124	CTGCGCGGTGAACAAACCGGAGGAAGTGGGGATGACATCAATCCCTCAATGAGG	1183
Db	758	CTGCGCGGTGAATAGCCGGAGGAAGTGAAGATGAAGTCAATCAATGAGCCCTTAATGCC	817
Qy	1184	TAGGGCTTACACGCTTAATACATGGCGGTGACAGAGGTTTCCCAACCCGCAAGGAGGAGC	1243
Db	818	CTGGGCGACACACGTGTCTACATATGGCCGGACAAAGGATCCGCAAGGAGGTAGAC	877
Qy	1244	TAACTCTGAAAGCGCGTGTATGTCGGATCGAGTCTGCAACTCGATCTCCGTGAAGTCG	1303
Db	878	TAACTCTCAAAAACCGTCTCAATGTCGGAATTTGACAGGCTTGACATCTGCTCATGAACCG	937
Qy	1304	GAATCGCTAGTAATCGCGATCAAC	1328
Db	938	GAATCGCTAGTAATCGCGATCAAG	962

RESULT 24	LOCUS	DEFINITION	ACCESSION
AQ957362	746 bp	DNA linear	GSS 28-JAN-2000
		LERAP67F LBRA Arabidopsis thaliana genomic clone LERAP6, genomic survey sequence.	
			AQ957362

VERSION	AQ957362.1	GI:6785063
KEYWORDS	GSS.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Vitridiplantae; Streptophyta; Embryophyta; Tracheophyta Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons rosids; eurosoids II; Brassicales; Brassicaceae; Arabidopsi-	
AUTHORS	Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utechtach, T. 1 (bases 1 to 746)	
TITLE	Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms	
JOURNAL	Unpublished (2000)	
COMMENT	Contact: Xiaoying Lin The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel.: 301 838 0200 Fax: 301 838 0208 Email: alie@tigr.org For additional information, see http://www.tigr.org/tcdb/ac/ac.html Seq primer: TP Class: shotgun.	
FEATURES	Location/Qualifiers	
source	1..746 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /ecotype="Landsberg erecta" /db_xref="taxon:3702" /clone="LERAP36" /clone_lib="LERA" /note="Organ: Leaf; Vector: pHOS1: Total genomic DNA was sheared to 0.9-1 kbp before ligation."	
ORIGIN		
Query Match	38.7%; Score 563.2; DB 9; Length 746;	
Best Local Similarity	86.3%; Pred. No. 3,1e-163;	
Matches 622; Conservative	0; Mismatches 99; Indels 0; Gaps	
Dy	85 AGTAATGACATCGGAAGCATTCAGAAGAGGGGGGTAAAGCATCGAATCGAAGATGCTTATAC	14
Dy	1 AGTAATATATCGGAACGTCTTTATATGGGGATTAATACTAGTCGAAGATTAGCTAATAC	60
Dy	145 CGATATTAATCTTAAGAGAAGACAGGGGATCGAAGACTTGCGCTTTGGAGCGCG	20
Dy	61 CGATATCGCCCTGAAGGGAGGAAGTAGGGGATCTTCGACCTTAAGCTTAAGAGAGGCGC	120
Dy	205 ATGTCATTAATCAATAGTGTGGGGTAAAGGCTTACCAAGGCGAGATCAATAGTGTGC	26
Dy	121 ATRTCTGATTAAGTAACTTGTGGGGTAAATGCTTACCAAGCTTGATCAATAGTGTGC	180
Dy	265 TGAGAGAGCACACGACCACTGGAGCTGAGACAGGCCCAAGACTCTTAGCGGAGGCGC	32
Dy	181 TGAGAGAGCACACGACCACTGGAGCTGAGACAGGCCCAAGACTCTTAGCGGAGGCGC	240
Dy	325 AGTGGGGAATTTTGACAATGGGCGCAAGCTGATCCAGCAATGCCCGGTGATGTAAGAA	38
Dy	241 AGTGGGGAATTTTGACAATGGGCGCAAGCTGATCCAGCAATGCCCGGTGATGTAAGAA	300
Dy	385 GGCCTTCGGGTTGTAAGCTCTTTCAGTCCAGAACAAAAGGTTCGTAATTAATCGTA	44
Dy	301 GGCCTTCGGGTTGTAAGCTCTTTCAGTCCAGAACAAAAGGTTCGTAATTAATCGTA	360
Dy	445 CCCATGACGGTATTCGACAGAAAGACACCGGCTAATCACTGTCGACGACCGCGTAATA	50
Dy	361 AGGCTACGGTATTCGACAGAAAGACACCGGCTAATCACTGTCGACGACCGCGTAATA	420
Dy	505 CGTAGGGTSCAAGCTTAATCGGAATTAATCGGGCTTAAGAGGTGCGCAGCGCGCTTTTA	56
Dy	421 CGTAGGGTSCAAGCTTAATCGGAATTAATCGGGCTTAAGAGGTGCGCAGCGCGCTTTGA	480
Dy	565 AGTCAGATGTGAATTCGCCGGCTTAATCGTGGGAATTCGTTGAAATCAACAAAGCTGA	62

Db 481 AGTCAGATGTGAATCCCGAGCTCAACTTGGGAACCTGCTTTGAAATCCGACACTAGA 540  
Qy 625 GTGTGACAGAGGAGGTGGAATTTCCATGTAGTCAAGTGAATGCGTAGAGATATGAGAGA 684  
Db 541 ATATGTCAAGAGGGGGGTAGAAATTTCCACGTGTAGCACTGAATGCGTAGAGATGTGAGAGA 600  
Qy 685 ACATCGATGGCGAAGGACGCTCTGCGGTAACTGACGCTCAGTCAAGAAAGCGTGGG 744  
Db 601 ATACCAATGGCGAAGGACGCTCTGCGGTAAATATGAGCGTCAATGCAAGAAAGCGTGGG 660  
Qy 745 GAGCAAAACGAGATTAGATACCTGTGTAGTCCAGCGCTTAAAGATGTCAATGTTGGT 804  
Db 661 GAGCAAAACGAGATTAGATACCTGTGTAGTCCAGCGCTTAAAGATGTCAATGTTGGT 720  
Qy 805 G 805  
Db 721 G 721

RESULT 25  
LOCUS CN757363 860 bp mRNA linear EST 20-MAY-2004  
DEFINITION ID0AAALCE08RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAALCE08 5',  
mRNA sequence.  
ACCESSION CN757363  
VERSION CN757363.1 GI:47531286  
KEYWORDS EST  
SOURCE Acyrthosiphon pisum (pea aphid)  
ORGANISM Acyrthosiphon pisum  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;  
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.  
1 (bases 1 to 860)  
Hunter, W., Martinez-Torres, D., Rabhe, Y., Sabater-Munoz, B.,  
Sten, D., Tagu, D. and Wincker, P.  
An expressed sequence tags database for the pea aphid Acyrthosiphon  
pisum  
Unpublished (2004)  
JOURNAL Contact: D. Tagu  
COMMENT INRA Rennes  
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France  
Tel: +33-2.23.48.51.65  
Fax: +33-2.23.48.51.50  
Risk of contamination by bacterial sequences from obligatory  
(Buchnera) or facultative endosymbionts. These sequences were  
obtained in the frame of the International Consortium of Aphid  
Genomics in collaboration with Genoscope  
PCR Primers  
FORWARD: CAGGAAACAGCTATGAC  
plate: 1 row: E column: 8.  
Location/Qualifiers

FEATURES  
Source  
1. 860  
/organism="Acyrthosiphon pisum"  
/mol\_type="mRNA"  
/cultivar="developmentstage"  
/db\_xref="taxon:7029"  
/clone="ID0AAALCE08"  
/tissue\_type="whole insect"  
/dev\_stage="nymphs and adults (parthenogenetic females)"  
/lab\_host="XLI-Blue"  
/note="Vector: PBS-SK minus; Site 1: BclRI; Site 2: XhoI;  
Sample name: ID0AA ; Plant growth place: Department of  
Ecology & Evolutionary Biology, Princeton University ;  
Soil conditions: Soil ; Sowing date: 01/06/1999 ;  
Harvesting date: 01/06/1999 ; Stress date: no stress ;  
Description: Aphids inoculated on one-week old *Vicia faba*  
under non-sterile conditions. All parthenogenetic stages  
and both winged and wingless adults were collected for  
library construction. ; experimental condition: long  
photoperiod (16-hr light/8-hr dark at 18 C)"

ORIGIN

Query Match 38 6%; Score 562.4; DB 7; Length 860;  
Best Local Similarity 81.4%; Pred. No. 5.7e-162;  
Matches 700; Conservative 0; Mismatches 156; Indels 4; Gaps 4;

Qy 353 GCGTATCCAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTGTAAAGCTCTTCACT 412  
Db 1 GCGTATCCAGCTATGCGCGGTGAGTGAAGAGGCTTAAAGGTCTTCACT 60  
Qy 413 CGAAGAAAGGTTACGGTAAATATGTGACCCATGACGGTATCGACAGAAAGCAC 472  
Db 61 GGGAGGAAAAAATAAATACTAATATTTATTTGTGACGTTACCCGAGAAAGCAC 120  
Qy 473 CGGCTTAATCTGCGCAGAGCGCGGTATATCGATGAGGTGCAAGGTTATCGAATTA 532  
Db 121 CGGCTTAATCTGCGCAGAGCGCGGTATATCGAAGGTGCAAGGTTATCGAATTA 180  
Qy 533 CTGGGCGTAAAGGTTGCGCAGAGCGCGCTTGTATGATGATGAAATCCCGGCTTAA 592  
Db 181 CTGGGCGTAAAGGCGCGGTATGATGATGATGATGATGATGATGATGATGATGAT 240  
Qy 593 CTGGGAAATGCGTTTGAATCTAAGAGTGTGCGAGGAGGAGTGAATCCCTGATG 652  
Db 241 CTAGGAATCTGATTTGAATCTGAAATCTAGAGTTCTGAGAGGAGTGAATCTTAG 300  
Qy 653 TGTAGAGTGAATGCGTGAATATGAGAAACATCGATGGCGAAGGAGCGCTCTGAG 712  
Db 301 TGTAGCGTGAATGCGTGAATATGAGAAATACCGTGGGAAAGGCGCTCTTAA 360  
Qy 713 TTAACATGACGCTCATGACAGAAAGCGGGAGCAACAGAGTTAGATACCTGTAG 772  
Db 361 CGAATCTGACTGAGGCGCGCAAGCGTGGGAGCAACAGAGTTAGATACCTGTAG 420  
Qy 773 TCCAGCGCTTAAACGATGTCAACTGATGTTGGGCTTATAGGCTTG-CTAAGAACT 831  
Db 421 TCCATGCGGTAAACGATGTCAACTGATGTTGGTTCAGAGAGTGAATCTTCCAGACT 480  
Qy 832 AACGCGTGAATGACCGCTGCGGAGTACGGTGGCAAGTTAACTCAAGAAATGA 891  
Db 481 AACGCAATTAAGTGAACCGCTGCGGAGTACGGTGGCAAGTTAACTCAAGAAATGA 540  
Qy 892 CGGGGACCGGCAACAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAAT 951  
Db 541 CGGGGACCGGCAACAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAAT 600  
Qy 952 CTTACCTTGAACATGTAGGGAATTTCTAGAGTATGATG-CTTGGGAAAGCTTAC 1010  
Db 601 CTTGATCTTGAACATGTAGGGAATTTCTAGAGTATGATG-CTTGGGAAAGCTTAC 660  
Qy 1011 ACAAGTGTGATGAGCTGTGTGATGCTGTGTGTGATGATGATGATGATGATGAT 1070  
Db 661 ACAAGTGTGATGAGCTGTGTGATGCTGTGTGTGATGATGATGATGATGATGAT 720  
Qy 1071 GAGCGCAACCTTGTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1129  
Db 721 GAGCGCAACCTTGTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 780  
Qy 1130 GTGACAAACCGAGGAAGGTGGGATGACCTCAAGTCTCATGG-CCCTTATGGGATG 1188  
Db 781 GTTATTAACCGAGGAAGGTGGGATGACCTCAAGTCTCATGG-CCCTTATGGGATG 840  
Qy 1189 CTTTACACGTTAATCAATGG 1208  
Db 841 CTTACACGCGTCAATATGG 860

RESULT 26  
LOCUS BZ464837 947 bp DNA linear GSS 13-DEC-2002  
DEFINITION BOKP76TR BO 1.6.2 KB. loc Brassica oleracea genomic clone BOKP76.  
genomic survey sequence.  
ACCESSION BZ464837  
VERSION BZ464837.1 GI:26752194  
KEYWORDS GSS.

SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 947)  
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Utebäck,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
TITLE Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
JOURNAL Genome Res. 15 (4), 487-495 (2005)  
PUBMED 15805490  
COMMENT Other GSSs: BONK15TR  
Contact: Chris Town  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Classes: sheared ends.  
FEATURES  
Location/Qualifiers  
source 1..947  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO100DH3"  
/db\_xref="taxon:3712"  
/clone="BONK15"  
/clone\_1lb="BO\_1.6.2\_KB\_tot"  
/note="Vector: pHOSt1; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pHOSt1 using BstXI linkers"  
ORIGIN  
Query Match 38.6%; Score 562; DB 9; Length 947;  
Best Local Similarity 76.9%; Pred. No. 7.8e-162;  
Matches 735; Conservative 0; Mismatches 195; Indels 26; Gaps 3;  
QY 166 AGCAGGGATCGAAAGACCTTGGCGCTTTGGAGCGCGCATCTGATTAGTCTAGTTGGT 225  
DB 6 AGCTAGAGGAGCAAGAGAGATCCGCCGAGGAGGAGGCTCGCTGATTAGTTAGTTGGT 65  
QY 226 GGGGTAAGGCTTACCAAGGCGAGCATGATGTTGTTGAGAGAGACAGACGACAC 285  
DB 66 GAGGCAATAGCTTACCAAGGCGATATGATGTTGTTGAGAGAGATGATGACACAC 125  
QY 286 TGGGACTAGACACGCGCCAGACTCTTACGAGGCGAGCATGTTGGGAAATTTTGGCAATG 345  
DB 126 TGGGACTAGACACGCGCCAGACTCTTACGAGGCGAGCATGTTGGGAAATTTTGGCAATG 185  
QY 346 GGGCGAAGCTGATCCAGCATGCGCGGTGAGTGAAGAGGCTTTCGGGTTGTAAGCTC 405  
DB 186 GGGCGAAGCTGATCCAGCATGCGCGGTGAGTGAAGAGGCTTTCGGGTTGTAAGCTC 245  
QY 406 TTTGAGTGAGAGAAAGGTTAAGTTAATGTGACCCGATGACGGTATCGACAGA 465  
DB 246 TTTTCCAGAGAGA-----AGCATATACGGTATCTGGGAAA 282  
QY 466 GAAGCAGCGGTAACTACGTGCGACAGCGCGGTAAATACGTAGGGTGAACCGTTAATC 525  
DB 283 TAAAGATCGGGTAACTCTGTGCGACAGCGCGGTAAATACGTAGGGTGAACCGTTAATC 342  
QY 526 GGAATTACTGGGCGTAAAGGGTGGCGAGCGCGCTTTGTAAGTCAAGTGAATCCCGG 585  
DB 343 GGAATTACTGGGCGTAAAGGGTGGCGAGCGCGCTTTGTAAGTCAAGTGAATCCCGG 402  
QY 586 GCTTAACCTGGGAATTGGTTGAAACTAACAAGCTAGAGTGTGCGAGGGAGGTGAA 645  
DB 403 GCTTAACCTGGGAATTGGTTGAAACTAACAAGCTAGAGTGTGCGAGGGAGGTGAA 462  
QY 646 TTTCATGTGTAGCAAGTGAATGCGTAGAGATGTGAAGAACATCGATGGCGAAGGACACC 705  
DB 463 TTTCGCGTGGAGCGGTGAATGCGTAGAGATGTGAAGAACATCGAAGCGGAAAGCATCTC 522

QY 706 TCCGTGGTTAACTGACGACTCATGACAGAAAGGTGGGAGCAACAGATTAATGATACC 765  
DB 523 TCGTGGGCGGACACTGACATGAGAGAGAAAGCTAGGGGAGCGAATGGATTGATACC 582  
QY 766 CTGTAGTGCACGCGCTTAAAGATGCAATGATGTTGGGCTTATTAAGGCTTG--GTA 823  
DB 583 CCAAGTATGCTTAAGCGGTAAAGATGATGATAGGGGCTGTGCGTATGACCGCGTCAATG 642  
QY 824 ACGAAGTAAAGCGGTGAAGTTGACCGCTTGGGAGTACGTTGCGAAGTTAAACTCAA 883  
DB 643 CTGTAGCTAAGCGGTTAAGTATCCGCTGGGAGTACGTTGCGAAGTAACTCAA 702  
QY 884 GGAATTGACGGGAGCCCGCAAGAGCGGTGATTAATGATTAATTCATGACAGCGAA 943  
DB 703 GGAATTGACGGGAGCCCGCAAGAGCGGTGATTAATGATTAATTCATGACAGCGAA 762  
QY 944 AAACCTTACCTACCTTGAATGATGAGCAATTTTCTAGAGATGATGATG--CTTGGGA 1002  
DB 763 GAACCTTACCAAGGCTTGAATGATGAGCAATTTTCTAGAGATGATGATG--CTTGGGA 822  
QY 1003 ACGCTAACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1062  
DB 823 ACGGAGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 882  
QY 1063 CCGGAGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1118  
DB 883 CCGGAGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 938  
RESULT 27  
B2461852/c 901 bp DNA linear GSS 13-DEC-2002  
LOCUS B2461852/c 901 bp DNA linear GSS 13-DEC-2002  
DEFINITION BONK15TR BO\_1.6.2\_KB\_tot Brassica oleracea genomic clone BONK15,  
genomic survey sequence.  
ACCESSION B2461852  
VERSION B2461852.1 GI:26745927  
KEYWORDS GSS  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 901)  
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Utebäck,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
TITLE Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
JOURNAL Genome Res. 15 (4), 487-495 (2005)  
PUBMED 15805490  
COMMENT Other GSSs: BONK15TR  
Contact: Chris Town  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Classes: sheared ends.  
FEATURES  
Location/Qualifiers  
source 1..901  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO100DH3"  
/db\_xref="taxon:3712"  
/clone="BONK15"  
/clone\_1lb="BO\_1.6.2\_KB\_tot"  
/note="Vector: pHOSt1; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pHOSt1 using BstXI linkers"  
ORIGIN  
Query Match 38.4%; Score 560.2; DB 9; Length 901;

Best Local Similarity 78.5%; Pred. No. 2.8e-161; Matches 696; Conservative 0; Mismatches 188; Indels 3; Gaps 2;

444 ACCATGACGGTATCGACAGAAAGGACCGGCTAACTAGTGGCAGACCGCGGTAAT 503  
 887 AGCAATACGGTATCTGGGAAATAGATCGGCTACTCTGTGCGACACCGCGGTAAT 828  
 504 ACCTAGGGGTCAAGCCGTTAATCGGAATTAATCTGGGCGTAAAGGTTGCCAGCGGCTTTGT 563  
 827 ACAGAGATCGAAGCGTTATCCGGAATGATGGCGTAAAGCGTCTAGTGGCTTTT 768  
 564 AAGTCAGATGGAATCCCCGGGCTTAACCTGGGAATTTGGTTGAAACTAACAAGCTAG 623  
 767 AAGTCCCGCTCAATCCAGGGCTCAACCTTGACAGCGGCTGAAACTCAAGCTTG 708  
 624 AGTGTGACAGAGGAGGTGGAATTCATGTGTACAGTGAATCCGTAGATATGGAAG 683  
 707 AGTACGTTAGGGGACAGAGGAAATTTCCGGTGGAGCGGTGAATGCTTAGATCGGAAG 648  
 684 AACATCGATGGCGAAGGACGCTCTCGGTTTAACTGACGCTCATGCAAGAAAGCGTGG 743  
 647 AACACCAACGGCGAAGACACTCTGCTGGCCGACACTGACACTGAGAGCAAGAAAGCTAGG 588  
 744 GAGCAAAACAGGATTAATGATACCTGTGTATGTCACGCGCTAAACGATGCACTAGTTGT 803  
 587 GAGCGAATGGATTAATGATACCTGATGCTTACCGCTTAAACGATGATATAGGCGCT 528  
 804 GGGCTTAATTAAGGCTTG--GTAAAGAAAGCTAAACGCTGAAGTTAAGCGGCTGGGAGTAC 861  
 527 GTGCGTATCGACCGGTGACGTGTGTAAGTAAACCGTTAATGATATCCGCTCGGAGATAC 468  
 862 GGTGCAAGATTAATACTCAAGAAATTTGACGGGACCCGCAACAGCGGTGATTAATGTG 921  
 467 GTTCGCAAGATGAAATCAAGAAATTTGACGGGCGCCGCAACAGCGGTGACATGTG 408  
 922 GATTAAATTCAGTCAACGCGAAACCTTACCTTACCTTGAATGATGACGAATTTTCTAG 981  
 407 GTTAAATTCAGTCAACGCGAAACCTTACCTTACCTTGAATGATGACCGGATCTCTTG 348  
 982 AGATAGATTAATG--CTTCCGGAAGGCTTAACAGGCTGATGATGCTGTGCTGACGTCT 1040  
 347 AAAAGAGAGGGGTGCTTGGGAAACGCGGACAGAGGTGATGATGCTGTGCTGACGTCT 288  
 1041 GTCTGAGATGTTGGGTTAATGCTCCGCAACGAGCGCAACCTTGTATTAATTTGCATCA 1100  
 287 GCCCTAAGGTGTGGGTTAATGCTCCGCAACGAGCGCAACCTTGTATTAATTTGCATCA 228  
 1101 TTTGGTGGGCACTTTAATGAGACTGCGGTGACAAACCGAGAGAGGTGGGATGACGT 1160  
 227 TTGAGTTTGAACCTTGAAACAGACTGCGGTGATTAAGCCGAGAGAGGTGAGATGACGT 168  
 1161 CAATGCTCATGCGCTTAAATGAGGCTTCAACGTAATCAATGCGGGGTACAGAG 1220  
 167 CAATGCTCATGCGCTTAAATGAGGCTTCAACGTAATCAATGCGGGGTACAGAG 108  
 1221 GTTGCAGAACCGCGAGGGGAGCTAATCTCAGAAAGCGGTGATGCTCGGATCGGAGTTC 1280  
 107 GTCCGATTCGCGAGGGGTGAGCTAATCCAAAACCGGTCTCAGTTCCGATTTGACAGC 48  
 1281 TGCAATCTGACTCCGTGAAGTGGAAATGCTAGTAATGCGGATGAG 1327  
 47 TGCAATCTGACTCCGTGAAGTGGAAATGCTAGTAATGCGGATGAG 1

RESULT 28  
 B2469058 1106 bp DNA linear GSS 13-DEC-2002  
 LOCUS B2469058/1.6.2 KB tot Brassica oleracea genomic clone BONH54,  
 DEFINITION genomic survey sequence.  
 ACCESSION B2469058  
 VERSION B2469058.1 GI:2674665  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea

## ORGANISM

Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

## REFERENCE

1 (bases 1 to 1106)  
 Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
 Utterback,T.R., Mortimer,J.R., White,O.R. and Town,C.D.

## TITLE

Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis

## JOURNAL

Genome Res. 15 (4), 487-495 (2005)  
 PUBMED  
 15805490

## COMMENT

Other GSSs: BONH54TR

Contact: Chris Town

7912 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..1106

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="T01000DH3"

/db\_xref="taxon:3712"

/clone="BONH54"

/note="Vector: pHD51; Site 1: BstXI; 1.6-2 kb sheared

/total DNA inserted into pHD51 using BstXI linkers"

## ORIGIN

## Query Match

Best Local Similarity 74.9%; Pred. No. 1.4e-160;

Matches 813; Conservative 0; Mismatches 220; Indels 53; Gaps 7;

3 TGAACGTGGCGGATCTTTTACATGCAAGTGAACCGGACGACGATGCTTCACT 62  
 1057 TGACCGCTGGCGGATCTTTTACATGCAAGTGAACCGGACGATGCTTCACT 1010  
 63 GGTGGCAGTGGCGGACGCGGTGAGTATGATCGGAAGCTATCCAGAAAGGCGGTAAC 122  
 1009 -GTTTCAAGTGGCGGACGCGGTGAGTATGATCGGAAGCTATCCAGAAAGGCGGTAAC 952  
 123 GCATGGAAGATGCTTATATCCGATATATCTTAAAGAGAAACGAGGATGAAAGA 182  
 951 AGCTGAAAGCGGCTGATATATCCCGTATGCTG--AGGACAAAGAGGATC----- 900  
 183 CTTTGGCTTTTGGAGCGGCGGATGCTGATTAAGTATGTTGGTGGGTTAAAGCTTACA 242  
 899 -----CGCCGAGGAGGGGCTCGGCTGATTAAGTATGTTGGTGGGTTAAAGCTTACA 845  
 243 AGGCGAGATGATGATGTTGCTGAGAGAGCAACAGCAACTGGGACTGAGACAGGC 302  
 844 AGGCGAGATGATGATGTTGCTGAGAGAGCAACAGCAACTGGGACTGAGACAGGC 785  
 303 CCAGATCTTACGAGGAGCAACAGTGGGAAATTTTGAACAATGGGCGCAACCTGATCCA 362  
 784 CCAATCTCTTACGAGGAGCAACAGTGGGAAATTTTCCGCAATGGGCGCAACCTGACGGA 725  
 363 GCAATGCCCGGTGATGAGAGGCTTCCGGTGTAAAGCTCTTTCACTGAGTGAAGAA 422  
 724 GCAATGCCCGGTGATGAGAGGCTTCCGGTGTAAAGCTCTTTCCGAGAAAG-- 667  
 423 AGTTACGGTAAATATCGTGAACCATGACGGTATCGACAGAAAGACCGGCTTAATA 482  
 666 -----AGCAATGACGGTATCTTGGGAAATGACATGCGGCTTAATC 628  
 483 GTGCGACAGACCGCGGTAATATCTAGGTGCAAGCTTAAATCGAAATTAATGAGGCGGTA 542  
 627 TGTGCGACAGACCGCGGTAATATCTAGGTGCAAGCTTAAATCGAAATTAATGAGGCGGTA 568

QY 543 AGGTCGCGAGGGGCTTGTAGTCAAGTGAATCCCGGGCTTAACCTGGGAATTG 602  
 DB 567 AGGTCGTAGTGGGCTTTTAAAGTCGCGCGTCAAAATCCAGGGCTCAACCTGGACAG 508  
 QY 603 CGTTGAAACTACAAAGCTAGAGTGTGCGAGAGGAGTGGAAATTCATGTGTAGCAGT 662  
 DB 507 CGGTGAAACTACAAAGCTTGAAGTGTGAGGAGAGGAAATTTCCGTGAGCGGTG 448  
 QY 663 AAATGCGTAGAGTATGAAAGAAATCATGATGCGAAAGGCACTCTGGGTAACTGA 722  
 DB 447 AAATGCGTAGAGTATGAAAGAAACCAACGCGAAAGCACTCTGGGTGCGACACTGA 388  
 QY 723 CGTCAATGACGAAAGCGTGGGAGCAACAGATTAAGATACCTGGTGTGTCACGCCCT 782  
 DB 387 CACTGAGAGACGAAACCTAGAGGAGGAGCAATGGAATTAAGATACCCAGTGTCTAGCCGT 328  
 QY 783 AAACGATGTCATAGTGTGTTGGGCTTATTAGGCTTG--GTAAACGAACTAAACGGCTGA 840  
 DB 327 AAACGATGATTAATAGGCGCTGTGCGTATGACCCCGTGCAGTGTCTGTAGCTAAACGGCTTA 268  
 QY 841 AGTTGACCGCTGGGAGATACGGTCCGAAAGATTAAATCTAAAGAAATGACCGGGAACC 900  
 DB 267 AGTATCCCGCTGGGAGATACGTTCCGAAAGAAATGAACTCAAGAAATGACCGGGGAGCC 208  
 QY 901 GCACAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAACCTTAACCTACCCCTT 960  
 DB 207 GCACAGCGGTGAGATGATGATGATTAATTCGATGCAACGCGAAACCTTAACCGAGGCTT 148  
 QY 961 GACATGAGCGAATTTTCTAGAGATAGATTAGT--CTTCGGGAGCGCTAACACAGGTGCT 1019  
 DB 147 GACATGCGGATCTCTTGAAGAGAGGGGTGCTTCGGGAGCGGACACAGGTGCT 88  
 QY 1020 GCATGCTGTGCTGACGCTGTGCTGAGATGTTGGTTAAGTCCGCAACGAGCGCAAC 1079  
 DB 87 GCATGCTGTGCTGACGCTGTGCTGAGATGTTGGTTAAGTCCGCAACGAGCGCAAC 28  
 QY 1080 CCTGT 1085  
 DB 27 CCTGT 22

RESULT 29  
 BZ474941/c 1069 bp DNA linear GSS 13-DEC-2002  
 LOCUS BZ474941  
 DEFINITION BONKR59tr BO\_1.6\_2 KB loc Brassica oleracea genomic clone BONKR59,  
 genomic survey sequence.  
 ACCESSION BZ474941  
 VERSION BZ474941.1 GI:26776372  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eustosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 1069)  
 Aylele,M., Haas,B.U., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
 Uteback,T.R., Wortman,J.R., White,O.R. and Town,C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 15805490  
 COMMENT Other\_GSSs: BONKR59TF  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.

FEATURES  
 source 1..1069

ORIGIN  
 Query Match 38.3%; Score 557.8; DB 9; Length 1069;  
 Best Local Similarity 74.3%; Pred. No. 1.6e-160;  
 Matches 806; Conservative 0; Mismatches 227; Indels 52; Gaps 6;  
 3 TGAAGCTGCGCGCATGCTTTACATGCAATGCAACGGCAGCAACGATGCTGCATCT 62  
 DB 1036 TGAAGCTGCGCGCATGCTTTACATGCAATGCAACGGCAGCAACGATGCTGCATCT 990  
 63 GGTGGCAGTGGCGGACGCGGTGATGATGATCGGAACGTATCCAGAAAGGGGGTAAAC 122  
 DB 989 TGTTCACATGCGCGGACGCGGTGATGATGATCGGAACGTATCCAGAAACGCGG 930  
 QY 123 GCATGAAAGATGTCTAATACCGCATATCTCTAAGAGAAAGCAGGGATCGAAAGA 182  
 DB 929 AGCTGAAACGCGCTGCTAATACCCGTAGGCTG--AGAGCAAAAGAGGAATC----- 878  
 QY 183 CCTTGCGCTTTGAGACCGCCGATGCTGATTAAGTGTGTTGGGTAAAGGCTACCA 242  
 DB 877 -----CGCCGAGAGAGGGCTGCGCTGATTAAGTGTGAGCAATAGCTTACCA 823  
 QY 243 AGCGCAGATCATGATGATGTTGCTGAGAGAGCAGCAGCAGCTGAGACTGAGACAGCG 302  
 DB 822 AGCGCATGATCATGATGCTGTGCGAGAGATGATCAGCAGCTGAGACTGAGACAGCG 763  
 QY 303 CCAGACTCTTACGAGAGAGCAGCAGATGAGGAAATTTGACATAGGCGCAGCTGATCA 362  
 DB 762 CCAGACTCTTACGAGAGAGCAGCAGATGAGGAAATTTGACATAGGCGCAGCTGACGA 703  
 QY 363 GCATGCGCGGTGAGTGAAGAGCGCTTGGGTTGTAAGCTCTTCACTGAGAAAGAA 422  
 DB 702 GCATGCGCGGTGAGTGAAGAGCGCTTGGGTTGTAAGCTCTTCACTGAGAAAGAA-- 645  
 QY 423 AGGTTACGGTAATATCGTGACCCATGACGGTATCGACAGAAAGACCGGCTTAACCTA 482  
 DB 644 -----AGCAATGACGATATCTGGGAAATAGATCGGCTTAACCT 606  
 QY 483 CGTGCCAGAGCGCGCGTAATACGTAGGTGCAAGCGTTAATTCGAATTAATCTGGGCTTA 542  
 DB 605 TGTGCCAGAGCGCGGTAAATACAGAGATGCAAGCGTTATCCGAAATGATTGGGCGTAA 546  
 QY 543 AGGTCGCGAGCGCGCTTGTAGTCAATGCAATGCCCGGGCTTAACCTGGGAATTG 602  
 DB 545 AGGTCGTAGTGGGCTTTTAAAGTCGCGCTCAATTCAGAGGCTCAACCTCGACAG 486  
 QY 603 CGTTGAAACTACAAAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGCAGT 662  
 DB 485 CGGTGAAACTACAAAGCTTGAAGTGTGAGGAGAGGAGAAATTTCCGTGAGCGGTG 426  
 QY 663 AAATGCGTAGAGTATGAAAGAAATCATGATGCGAAAGGCACTCTGGGTAACTGA 722  
 DB 425 AAATGCGTAGAGTATGAAAGAAACCAACGCGGAAAGCACTCTGGGTGCGACACTGA 366  
 QY 723 CGTCAATGACGAAAGCGTGGGAGCAACAGATTAAGTATACCTGTGTGTCCAGCCCT 782  
 DB 365 CACTGAGAGACGAAACCTAGAGGAGGAGGAAATGGAATTAAGTATACCTGTGTGTCCAGCCCT 306  
 QY 783 AAACGATGTCATAGTGTGTTGGGCTTATTAGGCTTG--GTAAACGAACTAAACGGCTGA 840  
 DB 305 AAACGATGATTAATAGGCGCTGTGCGTATGACCCCGTGCAGTGTCTGTACTAAACGGCTTA 246  
 QY 841 AGTTGACCGCTGGGAGATACGTCGCAAGATTAAATCTCAAGAAATGACCGGGAGCC 900  
 DB 245 AGTATCCCGCTGGGAGATACGTCGCAAGAAATGAAATCTCAAGAAATGACCGGGAGCC 186

/organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BONKR59"  
 /clone\_1b="BO\_1.6\_2 KB tot"  
 /note="Vector: pHO51; Site 1: Bact1; 1.6-2 kb sheared  
 total DNA inserted into pHO51 using Bact1 linkers"

Qy	901	GCACAACCGGTGATATTAATGCGATTAATTCGACACGCGAAGAAACCTTAACCAACCTT	960
Db	185	GCACAACGGGTGAGAGCATGTGGTTTAATTCATCATCAAGACGAAGAACCTTAACGAGGCTT	126
Qy	961	GACATGTAGCAATTTTCTAGAGATAGATTAGTG-CITTCGGAAACGCTAACACAGTGTCT	1019
Db	125	GACATGCCCGCAATTCCTTTGAAAGAGAGGGGTGCTCTTCGGGAACGGGAACAGAGTGGT	66
Qy	1020	GCATGGCTGTGTCAGCTCTGTGTGTCGTAGATGTTGGGTTAATGCCCGCAACGACGCCAAC	1079
Db	65	GCATGGCTGTGTCAGCTCTGTGTCGTAGATGTTGGGTTAATGCCCGCAACGACGCCAAC	6
Qy	1080	CCTTG 1084	
Db	5	CCTCG 1	
RESULT 30			
LOCUS	BZ437960	918 bp	DNA linear GSS 13-DEC-2002
DEFINITION	BONNV12TF BO.1.6.2_KB_tot Brassica oleracea genomic clone BONNV72,		
ACCESSION	BZ437960		
VERSION	BZ437960.1	GI:26692017	
KEYWORDS	GSS.		
SOURCE	Brassica oleracea		
ORGANISM	Brassica oleracea		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.		
AUTHORS	1 (bases 1 to 918) Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Uteerback,T.R., Wortman,J.R., White,O.R. and Town,C.D.		
TITLE	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res. 15 (4), 487-495 (2005)		
JOURNAL	15805490		
PUBMED	Other-GSS: BONNV72TR		
COMMENT	Contact: Chris Town		
	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA.		
	Tel: 301-838-3523		
	Fax: 301-838-0208		
	Email: cdtown@tigr.org		
	DNA is from a doubled haploid provided by Tom Osborn.		
	Seq primer: TF		
	Class: sheared ends.		
FEATURES	Location/Qualifiers		
source	1..918		
	/organism="Brassica oleracea"		
	/mol_type="genomic DNA"		
	/strain="TO100DH3"		
	/db_xref="taxon:3712"		
	/clone="BONNV72"		
	/clone_1ib="BO.1.6.2_KB_tot"		
	/note="Vector: pHOSt1, Site 1: BstXI; 1.6-2 kb sheared total DNA inserted into pHOSt1 using BstXI linkers"		
ORIGIN			
Query Match	38.0%;	Score 553.6;	DB 9; Length 918;
Best Local Similarity	78.8%;	Pred. No.36-159;	
Matches 688; Conservative	0;	Mismatches 184;	Indels 4; Gaps 3.
Qy	502	ATACGTAGGGTGCACGCGTTAATTCGGAATTACTGGCGCTTAAGGGTGGCGAGCGGCTT	561
Db	7	ATACGAGGATCCAGACCGTTATCCGGAATGATTTGGCGCTTAAGCGCTGTAGTGTGCTTT	66
Qy	562	GTAAGTCACATGTAATATCCCGGGGCTTAACCTGGGAATTTGGCTTTGAAATCAACAAGCT	621
Db	67	TTAAGTCCCGCTCAATCCAGGGCTCAACCTTGCAGCAGCGCGTGAATCAACAAAGT	126
Qy	622	AGAGTGTGCAGAGGAGGTGAATTCATGTGTGACAGTGAATTCGTAAGATATGGA	681

Db	127	TGAGTACGTAG333GAGAGGAAATTC	CGGTGAGCGGTGAATGCGT	AGATCGGAA	186
Qy	682	AGAAATCTGATGGCGAAGCAGCCTCCTGGGTTAA	CACTGACGCTCA	TGACGAAAGCGT	741
Db	187	AGAAACCAACCGCGAAGACA	CTCTGCTGGGCCGAC	CTGACACTGAGACGGAAC	246
Qy	742	GGGAGGCAACAGAGATTAGTAC	CCCTGGTAGTCCACGCCCTAA	ACATGTTCAC	801
Db	247	GGGAGGCAATGGGAGATTAGATAC	CCCCAGTAGTCTAGCCCTGAAACGAT	TGATATCTAGGCG	306
Qy	802	TTGGGCTTATTTAGGCTTG--GTAA	CAACTAAACGGGTAA	GGTGAACCGGCTGGGAGT	859
Db	307	CTGGCCCTATCGACCCCTGAGTCTGTATCA	ACGGTTAA	AGTATCCCGCTTGGGAGT	366
Qy	860	ACCGTCGCAAGATTAAACTCA	AAAGAAATTGACGGGAGCCCGAC	CAAGCGGTGATTTATG	919
Db	367	ACGTTCCGCAAGATGAAACTCA	AGAAATTGACGGGGGCCGAC	CAAGCGGTGAGCATG	426
Qy	920	TGATTTAATTCATGATGAC	CGCAAAACCTTACCTACCTTGACAT	GTAGCAATTTTCT	979
Db	427	TGTTTAAATTCATGATGAC	CAAGAAACCTTACCAAGGCTTGA	CATCCCGGAATCCTCT	486
Qy	980	AGAGATGATTATG--CTTGGGAA	CGGCTAACACAGGTGCGAT	TGGCTGTCAGCTC	1038
Db	487	TGAAGAAGAGGGGTGCTTCGGGAA	CGCGACACAGGTGTGATGGCTGTGTCAGCTC		546
Qy	1039	GTGTCGAGATATGTTGGGTTAA	GTCCCGCAACGACGCCCAACCTTGT	CATTTAATTGCCAT	1098
Db	547	GTGCCCTTAAGGTGTGGGTTAA	GTCCCGCAACGACGCCCAACCTCGT	TAGTTAGTTCAC	606
Qy	1099	CATTGGTTGGGACCTTTAAT	GAGATCTCCGGTGA	CAACCCGAGGAAGGTGGGATGAC	1158
Db	607	CGTTAGATTGGGAACCTCTGA	ACAGATCGCGGTATTAAGCCGAGGAAGGTGAGAGTAC		666
Qy	1159	GTCAAGTCCCATGSCCCTAT	TGGGTAGGGCTTGACAGTAA	TACATGGGCGGTACAGA	1218
Db	667	GTCAAGTATCATGCCCCCTAT	TGCGCTTATGGCGGAGACACAGTGTCTA	CAATGSCCGGACAA	726
Qy	1219	GGATTGCAACCCCGAGGGGAGCT	TAATCTCAGAAAGCGCGTGTAGTCCGAT	TCGGATTCGAG	1278
Db	727	GGGTCCGATCCCGAGGGGTAG	AGACTAACTCAAAAACCGGTCT	CAGTTGGATTGGAG	786
Qy	1279	TCTGCAACTGACTCCGTGA	ATGTCGAATGCTATGTAA	TGCGGATCAG--CATGTCGGG	1337
Db	787	GCTCAACTCGCCTGAC	TGAAGCCGGAATCGCTGTATGCGCGGT	CAGCCATACGGCGG	846
Qy	1338	TGAATAGCTTCCCGGGTCT	TGTACACACCGCGGTACACCAT	AGG	1383
Db	847	TGAATGCTTCCCGGGCTTGT	TACACACCGCCGTACACAT	TAGG	892

RESULT 31

Q0957363/c LOCUS

DEFINITION LERAP36TR LERA Arabidopsis thaliana genomic clone LERAP36, genomic survey sequence.

ACCESSION Q0957363

VERSION Q0957363.1 GI:6785064

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Uteerbach,T., Feldblum,T., Liang,F., Creasy,T. and Fraser,C.M. Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms

JOURNAL Unpublished (2000)

COMMENT Contact: Xiaoying Lin



The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: atc@tigr.org  
For additional information, see <http://www.tigr.org/cdb/at.html>  
Seq primer: TR  
Class: shotgun

## FEATURES

source

Location/Qualifiers  
1..716  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecoli\_type="landsberg erecta"  
/db\_xref="taxon:3702"  
/clone\_lib="LERA"  
/note="Organ: Leaf; Vector: pHOS1; Total genomic DNA was sheared to 0.9-1 Kbp before ligation."

## ORIGIN

Query Match 37.8%; Score 551.2; DB 9; Length 716;  
Best Local Similarity 86.7%; Pred. No. 1,6e-158;  
Matches 619; Conservative 0; Mismatches 93; Indels 2; Gaps 1;  
274 GACCAAGCAGCTGGAGCTGAGACAGCGCCAGACTCTTACGGAGAGGACAGTGGGGAA 333  
716 GACCAAGCAGCTGGAGCTGAGACAGCGCTCAGCTCTTACGGAGAGGACAGTGGGGAA 657  
334 TTTTGACATGGGCGCCAGAGCTGATTCAGCAATGCGCGTGAAGTGAAGAGGCTTCG 393  
656 TTTTGACATGGGCGCCAGAGCTGATTCAGCAATGCGCGTGAAGTGAAGAGGCTTCG 597  
394 GTTTGAAGCTCTTCACTGTCGAGAGAAAGTTAGCGTAATTAATCGTGAACCATGAG 453  
596 GTTGTAAAGCTCTTTCGCAAGGAGAAAGAACTTCTTAATTAAGAGTGAAGCTGAG 537  
454 GTATGACAGAAAGAACCGGCTAATCTAGTCGACAGCGCGGTAATAGTAGAGG 513  
536 GTATCTTGAATAAAGCAGCGGCTAATCTAGTCGACAGCGCGGTAATAGTAGAGG 477  
514 CAAGCTTAATCGAATTAATCTGCGGCTAAAGGGTGGCGAGCGCGCTTGAATGACAG 573  
476 CGAGCTTAATCGAATTAATCTGCGGCTAAAGGGTGGCGAGCGCGTGAATGACAG 417  
574 TGAATTCCTCGGCTTAATCTGCGGCTAAAGGGTGGCGAGCGCGCTTGAATGACAG 613  
416 TGAATTCCTCGGCTTAATCTGCGGCTAAAGGGTGGCGAGCGCGCTTGAATGACAG 357  
614 AGGAGTGGAAATTCATGTCAGTGAATGCGTAGAGTGAAGTGAAGTGAAGTGAAG 693  
356 AGGAGTGGAAATTCATGTCAGTGAATGCGTAGAGTGAAGTGAAGTGAAGTGAAG 297  
694 GCGAAGGAGCGCTCTGGGTTAACTGACGCTCATGCAAGAAAGCGTGGGAGCAACA 753  
296 GCGAAGGAGCGCTCTGGGTTAACTGACGCTCATGCAAGAAAGCGTGGGAGCAACA 237  
754 GGATTAGATACCTGCTGATGTCAGCGCTTAACGATGTCAGTGAATGTTGTTG -GGCTTA 811  
236 GGATTAGATACCTGCTGATGTCAGCGCTTAACGATGTCAGTGAATGTTGTTGTTG 177  
812 TTAAGCTTGGTAACGAGCTTAACGCGTGAAGTGAACCGCTGGGAGTACGCTGCAAG 871  
176 AATCAGTAGATACGAGCTTAACGCGTGAAGTGAACCGCTGGGAGTACGCTGCAAG 117  
872 TTTAACTCAAGAGATTGACGGGAGCCGACAGAGCGTGAATGATGATTAATTCG 931  
116 TTTAACTCAAGAGATTGACGGGAGCCGACAGAGCGTGAATGATGATTAATTCG 57  
932 ATGCAACGCGAAACCTTAACCTTGAACATGTAAGCAATTTCTAGAGAT 985  
56 ATGCAACGCGAAACCTTAACCTTGAACATGTAAGCAATTTCTAGAGAT 3

## RESULT 32

CV287759

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CV287759 915 bp mRNA linear EST 07-JUN-2005  
aofo1-13m83-h12 Aofo1 Asparagus officinalis cDNA clone  
aofo1-13m83-h12 5', mRNA sequence.  
CV287759  
CV287759.1 GI:52570310  
EST  
Asparagus officinalis (garden asparagus)  
Asparagus officinalis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;  
Asparagaceae; Asparagus.  
1 (bases 1 to 915)  
deRempfler, C., Carlson, J., Ma, H., Soltes, D., Soltes, P.,  
Oppehlmeier, D., Frohlich, M., Doyle, J., Tankley, S., Webb, M.,  
Leebens-Mack, J., Landherr, L., Illt, D., and Wall, K.  
Generation of ESTs from early male inflorescences of Asparagus  
officinalis  
Unpublished (2004)  
Contact: Claude deRempfler or James Leebens-Mack  
Mueller Laboratory  
Penn State University  
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn  
State University, University Park, PA 16802, USA  
Tel: 814 863 6413  
Fax: 814 865 9131  
Email: cws@psu.edu or jhl10@psu.edu  
The sequence provided is trimmed of vector and low quality regions.  
Full sequence and original trace file are available from the Plant  
Genome Network website (<http://pgn.cornell.edu>)  
Place: aofo1-13m83 row: h column: 12  
Seq primer: M13F.  
Location/Qualifiers  
1..915  
/organism="Asparagus officinalis"  
/mol\_type="mRNA"  
/db\_xref="PGN:aofo1-13m83-h12"  
/db\_xref="taxon:4686"  
/clone="aofo1-13m83-h12"  
/tissue\_type="male inflorescences"  
/lab\_host="SOLR"  
/clone\_lib="Aofo1"  
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;  
Site 2: XhoI; This is a directionally cloned,  
non-normalized library. This library has been generated by  
the Floral Genome Project (FGP). The Floral Genome Project  
is funded by NSF's Plant Genome Research Program  
(DBI-0115684). More information about the project can be  
obtained at <http://fgp.bio.psu.edu>"

## FEATURES

source

Query Match 37.6%; Score 548.2; DB 7; Length 915;  
Best Local Similarity 78.1%; Pred. No. 1,4e-157;  
Matches 684; Conservative 0; Mismatches 189; Indels 3; Gaps 2;  
444 ACCCATGACGCTATGACAGAAAGCAGCGGCTAATCTGTCAGCAGCGCGGCTAT 503  
40 ACCCATGACGCTATGACAGAAAGCAGCGGCTAATCTGTCAGCAGCGCGGCTAAG 99  
504 AGTGGGGGCAAGGGTTAATGAGAAATTAATGAGGAGTGAAGGAGGAGGAGGAGT 563  
100 AAGAGAGATGCAAGGCTTAATCGGAATGATTTGGGCTTAAGGCTCTGTAAGTGGCTTTTC 159  
564 AAGTGAATGTAATATCCCGGCTTAACCTTGAAGTTCGTTTGAATCAAAAGCTAG 623  
160 AAGTGGCGGCTCAATATCCAGGCTCAACCTTGAAGGAGGAGGAGGAGGAGGAGG 219  
624 AGTGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683  
220 AGTGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 279  
684 AACATGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 743



Db 280 AACACCAACGGCGAAAGCACTGCTGGGCCGACCTGACCTGAGAGAGAAAGCTAGG 339  
 Qy 744 GAGGCAAAACGAGTTAGTACCTGTAAGTCCAGCCCTTAAAGATGCACTAGTTGTT 803  
 Db 340 GAGGCAAAATGGATTAATATCCCAATGATCTTAAGCCGTAAAGATGATCTAGGCGCT 339  
 Qy 804 GGGCTTATTAAGGCTTG--GTAAAGAACTTAACGCGTGAAGTTGACCGCTGGGAGATAC 861  
 Db 400 GTGCGTATGACCCCGTGCAGTGTGATGCTTAAGCCGTAAATGATCCGCTGGGAGATAC 459  
 Qy 862 GTTCGCAAGATTAATACTCAAGGAATTAACGCGGAGCCGCAACAGGCTGGATTAATG 921  
 Db 460 GTTCGCAAGATTAATACTCAAGGAATTAACGCGGAGCCGCAACAGGCTGGATTAATG 519  
 Qy 922 GATTAATTCATGCAAGCGGAAACCTTACCTTACCTTGAATGATGCAATTTCTAG 981  
 Db 520 GTTAAATTCATGCAAGCGGAAACCTTACCGGCTTGAATGATGCAATTTCTAG 579  
 Qy 982 AGATAGATTAAGT-CTTCGGAACGCTTAACAGAGTCTGATGCTGTCAGCTCGT 1040  
 Db 580 AAAGAGAGGGGTGCTTCGGAACGCGGACAGAGTGTGATGCTGTCAGCTCGT 639  
 Qy 1041 GTGCTGATGATGTTGGTTAAGTCCGCAACGAGCGCAACCTTGTCAATTAATTCCTCA 1100  
 Db 640 GCGTAAAGTGTGGTTAAGTCCGCAACGAGCGCAACCTTGTGTAGTTGCTCCCA 699  
 Qy 1101 TTTGCTGGGCACTTAAATGAGTCCCGTGAACAAACCGGAGGAGTGGGATGAGT 1160  
 Db 700 TTGAGTGTGAACCTTAACAGACTGCGGTGAATGCGGAGGAGTGAATGAGT 759  
 Qy 1161 CAAGTCCCTCAATGAGGCTTATGAGGCTTCAACAGCTAATACATGAGCGGCTACAGAG 1220  
 Db 760 CAATGCTCATGATCCCTTATGCTGAGCGGACAGACAGTGTCAATGAGGCGGCAAG 819  
 Qy 1221 GTTCCCAACCGGAGGAGGAGCTAATCTCAAGAGCGGCTGTAATCGGATGAGT 1280  
 Db 820 GTGCGATATCCCGGAGGAGGAGCTAATCTCAAGAGCGGCTGTAATCGGATGAGT 879  
 Qy 1281 TGCACTCTGATCTCCGTGAAGTCCGAATCGCTAGTAA 1316  
 Db 880 TGCCATCTGCTGATGAAGCCGGAATCGCTAGTAA 915

RESULT 33  
 CV292617 875 bp mRNA linear EST 08-JUN-2005  
 LOCUS aof01-2ms4-a01 Aof01 Asparagus officinalis cDNA clone  
 DEFINITION aof01-2ms4-a01 5', mRNA sequence.  
 CV292617  
 ACCESSION CV292617.1 GI:52580066  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Asparagus officinalis (garden asparagus)  
 Asparagus officinalis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;  
 Asparagaceae; Asparagus.  
 1 (bases 1 to 875)  
 dePamphilis,C., Carlson,J., Ma,H., Solis,D., Solis,P.,  
 Openheimer,D., Frohlich,M., Doyle,J., Tanksley,S., Webb,M.,  
 Leebens-Mack,J., Landherr,L., Ilut,D. and Wall,K.  
 Generation of ESTs from early male inflorescences of Asparagus  
 officinalis (2004)  
 Unpublished  
 Contact: Claude dePamphilis or James Leebens-Mack  
 Penn State University  
 Mueller Laboratory  
 208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn  
 State University, University Park, PA 16802, USA  
 Tel: 814 863 6413  
 Fax: 814 865 9131  
 Email: cw43@psu.edu or jh110@psu.edu  
 The sequence provided is trimmed of vector and low quality regions.

Full sequence and original trace file are available from the Plant  
 Genome Network website (<http://pgn.cornell.edu>)  
 Plate: aof01-2ms4 row: a column: 01  
 Seq primer: M13F.  
 Location/Qualifiers  
 1..875  
 /organism="Asparagus officinalis"  
 /mol\_type="mRNA"  
 /db\_xref="PDB:aof01-2ms4-a01"  
 /db\_xref="taxon:4686"  
 /clone="aof01-2ms4-a01"  
 /issue\_type="male inflorescences"  
 /lab\_host="SOLR"  
 /clone\_lib="aof01"  
 /note="Vector: pBluescript SK (+/-); Site\_1: EcoRI;  
 Site\_2: XhoI; This is a directionally cloned,  
 non-normalized library. This library has been generated by  
 the Floral Genome Project (FGP). The Floral Genome Project  
 is funded by NSF's Plant Genome Research Program  
 (DBI-011684). More information about the project can be  
 obtained at <http://fgp.bio.psu.edu>"

ORIGIN  
 Query Match 37.2% Score 542.2; DB 7; Length 875;  
 Best Local Similarity 78.6% Pred. No. 1e-155;  
 Matches 686; Conservative 0; Mismatches 183; Indels 4; Gaps 3;

Qy 484 GTGCCAGACGCCGCGTAATACGTAAGGAGGCAAGCGTTAATCGGAATTAATCTGCGGCTTAA 543  
 Db 3 GTGCCAGACGCCGCGTAAGAGACAGAGATGCAAGCGTTAATCGGAATTAATCTGCGGCTTAA 62  
 Qy 544 GGGTGGCAGGCGGCTTGAATGATGATGAAATTCCTCGGCTTAACTCGGAATTCG 603  
 Db 63 GCGTGTAGTGGCTTTTCAGATTCGCGCTCAATTCGAGGCTCAACCTGGAACGGC 122  
 Qy 604 GTTTGAACCTCAAAAGCTAAGTGTGGCAGAGGAGGAGTTCATGTGTAGCACTGTA 663  
 Db 123 GTTGAACCTCAAAAGCTAAGTGTGGCAGAGGAGGAGTTCATGTGTAGCACTGTA 182  
 Qy 664 AATGCGTAGATATGAGAGCAATCGATGAGCGGAGCGAGCTCTCGGTTAACTGATC 723  
 Db 183 AATGCGTAGATATGAGAGCAATCGATGAGCGGAGCGAGCTCTCGGTTAACTGATC 242  
 Qy 724 GCTCATGACGAAAGCGTGGGAGCAACAGATTAATACCTGATGATCCAGCCCTTA 783  
 Db 243 ACTGAGAGAGAAAGCTAGGAGGAGCAAAATGAGATTAATACCTGATGATCCAGCCCTTA 302  
 Qy 784 AAGCATCTCACTAGTTGTTGGGCTTATTAAGGCTTG--GTAAAGAACTTAACGCGTGA 841  
 Db 303 AAGCATCTCACTAGTTGTTGGGCTTATTAAGGCTTG--GTAAAGAACTTAACGCGTGA 362  
 Qy 842 GTTACCGGCTGGGAGTACGCTGCGCAAGATTAACTCAAGGAATTAAGCGGAGCCG 901  
 Db 363 GTATCCGCGCTGGGAGTACGCTGCGCAAGATTAACTCAAGGAATTAAGCGGAGCCG 422  
 Qy 902 CACAAGGCTGATTAATGATTAATTCGATGCAAGCGGAAACCTTAACCTTACCTTG 961  
 Db 423 CACAAGGCTGATTAATGATTAATTCGATGCAAGCGGAAACCTTAACCTTACCTTG 482  
 Qy 962 ACATGAGCAATTTTCTAAGATTAATGATG--CTTCGGAACGCTTAACAGAGTCTG 1020  
 Db 483 ACATGAGCAATTTTCTAAGATTAATGATG--CTTCGGAACGCTTAACAGAGTCTG 542  
 Qy 1021 CATGCTGTGCTGCTGCTGCTGATGATGTTGGTTAAGTCCGCAACGAGCGCAAC 1080  
 Db 543 CATGCTGTGCTGCTGCTGCTGATGATGTTGGTTAAGTCCGCAACGAGCGCAAC 602  
 Qy 1081 CTTCATTAATTCATTAATTCGTTGGGCACTTTAATGATGATGCTGCGGTGACAAACCG 1140  
 Db 603 CTTCATTAATTCATTAATTCGTTGGGCACTTTAATGATGATGCTGCGGTGACAAACCG 662  
 Qy 1141 GAGGAGGTGGGATGAGTCAATCTCAATGCGCTTAAGGAGGCTTCAACAGCTAA 1200

Db 663 GAGGAAGTGAAGATGATGCTCAAGTATCATGCTCCCTTATGCTCGGGGACACACAGTCG 722

Qy 1201 TACAAATGCGCCCTTACAGAGGGTTGCCAAACCCGAGAGGGAGGCTAAATCTCAGAAAGCGG 1260

Db 723 TACAAATGCGCGGAGCAAAAGGGTCCGATCCCGAGAGGGTGAAGCTAACCTCCAAAACCCCGT 782

Qy 1261 TCGTAGTCGGATCGAGTCTGCAACTCGAATCCGTAATCGGTAATCGCTAGTAAATCGC 1320

Db 783 CTCATGTCGGATTCGACGCTGCAACTCGCTCAGTAAGAACCGGAATCGCTAATTAATCGC 842

Qy 1321 GGATCAG-CATGTCGCGGTGAATACGTTCCCGG 1352

Db 843 CGGTCAAGCCATACGCGCGTGAATTCGTTCCCGG 875

RESULT 34

LOCUS Bz279543 704 bp mRNA linear EST 09-APR-2002

DEFINITION Bz279543 Y. Ogihara unpublished cDNA library, Wh\_r Triticum aestivum cDNA clone whr3h15 5', mRNA sequence.

ACCESSION Bz279543

VERSION Bz279543.1 GI:20102041

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 704)

AUTHORS Ogihara, Y. and Murali, K.

TITLE Expressed genes in Triticum aestivum

JOURNAL Unpublished (2002)

COMMENT Contact: Tadao Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers

FEATURES

source

1..704

/organism="Triticum aestivum"

/mol\_type="mRNA"

/culivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="whr3h15"

/isue.type="root"

/dev\_stage="Peekes' scale 1"

/clone\_lib="Y. Ogihara unpublished cDNA library, Wh\_r"

ORIGIN

Query Match 37.2%; Score 541.6; DB 3; Length 704;  
Best Local Similarity 87.9%; Pred. No. 1.5e-155;  
Matches 603; Conservative 0; Mismatches 79; Indels 4; Gaps 1;

Qy 1 ATTGAACGCTGGGAGTATGCTTTACACATCAAGTCGACGAGACAGCATCTTCGAT 60

Db 23 ATTGAACGCTGGGAGTATGCTTTACACATCAAGTCGACGAGCGCGG---GGCAAC 78

Qy 61 CTGGTGGCGATGGCGGACGGGTGATGATCGGAACGATTCAGAAAGAGGGGGTGA 120

Db 79 CTGGCGCGAGTGGCCAGCGGTGATGATATATCGGAACGTAACCAAGATGGGGAT 138

Qy 121 ACGCATCGAAAGATGTGCTAATACCGCATATCTCTTAAGAGGAAAGACAGGGATGAAA 180

Db 139 ACGTACGAAAGATGTAGCTAATACCGCATATCTTAAGGATGAAAGTGGGATCGCAA 198

Qy 181 GACCTTGCGTTTGGAGCGCGCGATCTGATTAAGTATGCTGGGTAAGGCTTC 240

Db 199 GACCTATGCTCTGGAGCGCGCGATATGATTAAGTATGCTGGTAAAGGCTTC 258

Qy 241 CAAGGCGAGCATGATGATGCTGAGAGGACGACGACACACTGGGACTGAGACAG 300

Db 259 CAAGCGACGATCAGTACGTGCTGAGAGGACGACACGCACTGGGACTGACACAG 318

Qy 301 GCCCAAGCTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAAATGGGCGCAAGCTGATC 360

Db 319 GCCCAAGCTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAAATGGGCGCAAGCTGATC 378

Qy 361 CAGCAATGCGCGTGAATGAGAGAGGCGCTTCGGGTTGTAAGCTCTTTCACTGAGAAAG 420

Db 379 CAGCAATGCGCGTGAATGAGAGAGGCGCTTCGGGTTGTAAGCTCTTTTGTGAGGAAAG 438

Qy 421 AAGGTACGTTAAATTAATCTGACCATGACGATATGACAGAAAGACCGGCTAAC 480

Db 439 AAGCGATTTGCTTAATATCTGCTTAATGACGATCTGAAGAAATAGCACCGGCTAAC 498

Qy 481 TACGTCCAGCAGCCCGCGTATAGTGGGTCGAAGCGTTAATCGGAATTACTGGGCGT 540

Db 499 TACGTCCAGCAGCCCGCGTATAGTGGGTCGAAGCGTTAATCGGAATTACTGGGCGT 558

Qy 541 AAGGGTGGCAGCGCGCTTTGTAAGTCAGATGTGAAATCCCGGGCTTTAAGCTGGAAAT 600

Db 559 AAGCGTGGCAGCGCGCTTTGTAAGTCAGATGTGAAATCCCGGGCTTTAAGCTGGAAAT 618

Qy 601 TCGTTTGAACCTAACAAGCTTAGTGTGCGACAGAGGAGTGAATTCATGTGACAG 660

Db 619 TCGTTGAGACTGCAAGGCTTGAATCTGCGACAGAGGCGGTGAATTCACGTTAGCAG 678

Qy 661 TGAATGCTAGATATGAGAGAAC 686

Db 679 TGAATGCTAGATATGAGAGAAC 704

RESULT 35

LOCUS Bz502068/c 1096 bp DNA linear GSS 16-DEC-2002

DEFINITION BONIT25TR BO.1.6.2\_KB\_tot Brassica oleracea genomic clone BONI125, genomic survey sequence.

ACCESSION Bz502068

VERSION Bz502068.1 GI:27019377

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 1096)

AUTHORS Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utecher, T.R., Wortman, J.R., White, O.R. and Town, C.D.

TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis

JOURNAL Genome Res. 15 (4), 487-495 (2005)

PubMed 15805490

Other\_GSSs: BONI125TF

COMMENT Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdcow@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TP  
Class: sheared ends.  
Location/Qualifiers

FEATURES

source

1..1096

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="TO1000DH3"

/db\_xref="taxon:3712"

/clone="BONI125"

/clone\_lib="BO.1.6.2\_KB\_tot"

/note="vector: pHSI1; Site 1: BcXI; 1.6-2 kb sheared total DNA inserted into pHSI1 using BcXI linkers"

ORIGIN

Query Match 37.1%; Score 540.8; DB 9; Length 1096;  
 Best Local Similarity 77.2%; Pred. No. 2,9e-155;  
 Matches 696; Conservative 0; Mismatches 202; Indels 4; Gaps 3;

560 TTCTAAGTCAAGTGAATCCCGGCTTAACCTGGGAATTCGTTTGAATCAAAAG 619  
 1095 TTTTAAGTCCCGTCAAAATCCAGGGCTCAACCTGACAGAGCGGTGAAACTACAG 1036  
 620 CTAGAGTGTGACAGAGGAGTGAATTCATGTGTAGCAGTGAATGCTAGAGATATG 679  
 1035 CTGAGTACAGTGAAGGAGGAGGAAATTCGAGTGAAGCGGTGAATGCGTAGAGATCG 976  
 680 GAAAGATCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 739  
 975 AAGAGACACCAAGCGGCAAGAGCACTCTGCTGGGCGGCACTGACACTGAGAGAGCAAGC 916  
 740 GTGGGAGCAACAGATTGATACCTGATGTCAGCGCCCTAAAGATGTCACATGAT 799  
 915 TAGGGAGCAATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 856  
 800 TGTGGGCTTATTTAGGCTTG--GTAAAGAGCTAAGCGGTGAAGTTGACCGCTGGGGA 857  
 855 CGCTGTGCTATGACACCGGTGACAGTGTGATGATGATGATGATGATGATGATGATG 796  
 858 GTACGCTGCAAGATTAACTCAAGGAATTGACGGGAGACCGGCAAGAGCGGTGATTA 917  
 795 GTACGCTGCAAGATTAACTCAAGGAATTGACGGGAGACCGGCAAGAGCGGTGATTA 736  
 918 TGTGATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 977  
 735 TGTGATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 676  
 978 CTAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1036  
 675 CTGAG 616  
 1037 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1096  
 615 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 556  
 1097 ATCATTTGTTGGGCACTTTATGATGATGATGATGATGATGATGATGATGATGATG 1156  
 555 ACCGTTAGTTTGAACCTGGAACGATGCGGTGATTAAGCCGAGAGAGATGAGATG 496  
 1157 ACCTCAAGTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1216  
 495 ACCTCAAGTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 436  
 1217 GAGGAGTGGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1276  
 435 AAGGAGTGGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 376  
 1277 AGTGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1335  
 375 AGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 316  
 1336 GGTGAATGCTTCCCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATG 1395  
 315 GGTGAATGCTTCCCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATG 256  
 1396 CAGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1455  
 255 CAGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 196  
 1456 TG 1457  
 195 TG 194

RESULT 36  
 DN468747 DN468747 787 bp mRNA linear EST 01-AUG-2005  
 LOCUS  
 DEFINITION USDA-FP\_142729 Diaphorina citri Kuwiyama (Hemiptera: Psyllidae)

Diaphorina citri cDNA clone WHD040\_D02 5', mRNA sequence.  
 DN468747  
 DN468747.1 GI:71538204  
 EST.  
 Diaphorina citri  
 Diaphorina citri  
 Diaphorina citri  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Psyllidae; Psylliformes;  
 Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Psylliformes;  
 Psyllidae; Psyllidae; Diaphorina.  
 1 (bases 1 to 787)  
 Hunter, W.B., Dang, P.M. and McKenzie, C.L.  
 Gene expression in adult Asian citrus psylla, Diaphorina citri  
 Kuwiyama (Hemiptera: Psyllidae)  
 Unpublished (2006)  
 JOURNAL  
 COMMENT  
 Contact: Wayne Hunter, Phat Dang, USDA, ARS  
 U.S. Horticultural Research Lab, Subtropical Insect Research Unit  
 USDA-ARS  
 2001 South Rock Road, Ft. Pierce, FL 34945, USA  
 Tel: (772) 462-5898, (772) 462-5940  
 Fax: (772) 462-5986  
 Email: whunter@usnrl.ars.usda.gov  
 Seq primer: T3 primer.  
 Location/Qualifiers  
 1. 787

FEATURES  
 source  
 /organism="Diaphorina citri"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:121845"  
 /clone="WHD040\_D02"  
 /sex="Mixed genders"  
 /tissue\_type="whole body"  
 /dev\_stage="adults"  
 /lab\_note="X11-Blue"  
 /clone\_lib="Diaphorina citri Kuwiyama (Hemiptera: Psyllidae)"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; source: Colonies maintained by Dr. McKenzie, USHRL, Fort Pierce, FL. A high quality EST with at least 200 contiguous bases at Trace Reader score of 20 or better. Library construction by Laura Hunicutt, USDA, ARS, U.S. Horticultural Res. Lab, Ft. Pierce, FL, USA."

#### ORIGIN

Query Match 36.4%; Score 530.6; DB 8; Length 787;  
 Best Local Similarity 84.2%; Pred. No. 3.9e-152;  
 Matches 639; Conservative 0; Mismatches 104; Indels 16; Gaps 3;

1 ATTGAAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60  
 33 ATTGAAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 88  
 61 CTGCTGACGAGTGGCGGAACGGGTGATGATGATGATGATGATGATGATGATGATGATG 120  
 89 CTGCTGACGAGTGGCGGAACGGGTGATGATGATGATGATGATGATGATGATGATGATG 148  
 121 ACCGATGAAAGATGCTTAATACCGCAT-TACTCTAAGAGAGAGAGAGAGAGAGG 174  
 149 ACTAGTGAAGAAATTAATCTAATACCGCATCAAAATCTAAGAGAGAGAGAGAGAGG 208  
 175 -----TCGAAGACCTTGGCTTTTGAAGCGGCGGAGTGTCTGATTAAGTATGTTG 228  
 209 AGTATTTTAAAACTTAATGCTTTTGAAGCGGCGGAGTATCTGATTAAGTATGTTG 268  
 229 GTAAAGCTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 288  
 269 GTAAAGCTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 328  
 289 GACTGAGACACGCGCCGAGTCTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 348  
 329 GACTGAGACACGCGCCGAGTCTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 388  
 349 GCAAGCTGATCCAGCATGCGGCTGAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 408  
 389 GAAAGCTGATCCAGCATGCGGCTGAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 448

Qy	409	CAGTGCAGAAAGAAAAGTTACGGTAATAATATCTGACCCATGACGGGTATCGACAGAAGAA	468
Db	449	TATCAGGAAAGAAACGAATTAATATATATCATTTAATTTATATGACGGTACTGAAGAAATA	508
Qy	469	GCACCGGCTTAATCACTGCGCAGACGCCGGGGTAAATCGTAGGGTGCAGCGTTAATCGA	528
Db	509	GCACCGGCTTAATCACTGCGCAGACGCCGGGGTAAATCGTAGGGTGCAGGGTTAATCGG	568
Qy	529	ATTACTGGGCGTAAAGGGTGGCCAGCGGCGCTTTGAATGATCATATGTGAAATCCCCGGGCT	588
Db	569	ATTATTTGGGCGTAAAGCGTACGTAGGCGGCTTTTGTGAACAGATGTGAAATGCTTGGGCT	628
Qy	589	TTACCTGGGAATTTGGCTTTGAAACTCAAAAGCTAGAGTGTGGCAGAGGGAGGTGAATTC	648
Db	629	CACCTTGGAAATTTGCAATTTGTGACTCCAAAGCTAGAGTGTGTAGAGGGAGATGAATTC	688
Qy	649	CATGTGTACAGTGAATGCTGAGATATATGAAAGAACATCGATGGCGAAAGGAGCGCTCC	708
Db	689	CACATGTACAGTGAATGCTGAGATATATGTAAGAGAAATCCGATGGCGAAAGGAGTCTCC	748
Qy	709	TTGGTTAACTGACGCGCTCATGCAAGAAAGCTGGGGAG	747
Db	749	TGGGTTAACTGACGCGCTCATGCAAGAAAGCGTGGGGAG	787
RESULT 37			
LOCUS	BH527452	862 bp	DNA
DEFINITION	BGV36TR BGV Brassica oleracea genomic clone BGV36, genomic survey sequence.		
ACCESSION	BH527452		
VERSION	BH527452.1	GI:11735537	
KEYWORDS	GSS.		
SOURCE	Brassica oleracea		
ORGANISM	Brassica oleracea		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustosids II; Brassicales; Brassicaceae; Brassica.		
AUTHORS	1 (bases 1 to 862) Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Utecherback,T.R., Wortman,J.R., White,O.R. and Town,C.D.		
TITLE	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis		
JOURNAL	Genome Res. 15 (4), 487-495 (2005)		
PUBMED	15805490		
COMMENT	Other GSSs: BGV36TR Contact: Chris Town		
ORIGIN	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA.		
	Tel: 301-838-3523		
	Fax: 301-838-0208		
	Email: cdtown@tigr.org		
	DNA is from a doubled haploid provided by Tom Osborn.		
	Seq primer: TR		
	Class: sheared ends.		
FEATURES	Location/Qualifiers		
source	1..862		
	/organism="Brassica oleracea"		
	/mol_type="genomic DNA"		
	/strain="TO1000DH3"		
	/db_xref="taxon:3712"		
	/clone="BGVP36"		
	/clone_1lb="BGV"		
	/note="Vector: pHD1, Site 1: BstXI, 2-3 kb sheared genomic DNA inserted into pHD1 using BstXI linkers"		
Query Match	36.2%	Score 528; DB 9; Length 862;	
Best Local Similarity	77.8%	Prod. No. 2.5e-151;	
Matches 666; Conservative	0;	Mismatches 170; Indels 26; Gaps 3	
207	GTCTGATTAAGTAGTTGTTGGGGTAAAGGCTTCAAGGCGACGATCATGATTTGGTCTG	266	

[illegible]

REFERENCE 1 (bases 1 to 846)  
 AUTHORS Ayele,M., Haas,B.U., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
 Utterback,T.R., Mortman,J.R., White,O.R., and Town,C.D.  
 TITLE Whole genome shotgun sequencing of *Brassica oleracea* and its  
 application to gene discovery and annotation in Arabidopsis  
 JOURNAL Genome Res. 15 (4), 487-495 (2005)  
 PUBMED 15805490  
 COMMENT Other\_GSSs: BOGZ83TR  
 Contact: Chris Town  
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TP  
 Class: sheared ends.

FEATURES  
 source  
 1. 846  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO100DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOGZ83"  
 /note="Vector: pHOS1; Site\_1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN  
 Query Match 36.1%; Score 526.2; DB 9; Length 846;  
 Best Local Similarity 78.0%; Pred. No. 8.9e-151;  
 Matches 659; Conservative 0; Mismatches 183; Indels 3; Gaps 2;

487 CCAGCAGCCCGGATATACGAGGAGGTCAGACGCTTAATCGAATTACTGGGCGTTAAAGG 546  
 846 CCGACACCGCGGATATACGAGGAGGTCAGACGCTTAATCGAATTACTGGGCGTTAAAGG 787  
 547 TGCGCAGCGCGCTTGTAGTCAAGATGTAATCCCGGGCTTAACTGGGAATTCGCTT 606  
 786 TCTGTAGTGGCTTTTAAGTCCCGCTCAATCAGAGGCTCAACCTGGACAGCGCGGT 727  
 607 TGAACCTACAAAGCTAGAGTGTGCAAGGAGGAGTGCATTCATGTGTAGCATGAAAT 666  
 726 GGAACCTACAAAGCTAGAGTGTGCAAGGAGGAGGAGTGCATTCATGTGTAGCATGAAAT 667  
 667 GCGTAGAGATGGAAGACATCGATGCGGAGGAGGAGGAGTGCATTCATGTGTAGCATGAAAT 726  
 666 GCGTAGAGATGGAAGACATCGATGCGGAGGAGGAGGAGTGCATTCATGTGTAGCATGAAAT 607  
 727 CATGCAAGAAAGCTGGGAGGAGCAAGAGATTAGTACCTGTGATGTCACGCCCTTAAC 786  
 606 GAGAGAGAAAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 547  
 787 GATGTCAACTAGTGTGGGCTTATTAAGCTTG- GTACGAAGCTTAACCGCTGAAGTT 844  
 546 GATGTCAACTAGTGTGGGCTTATTAAGCTTG- GTACGAAGCTTAACCGCTGAAGTT 487  
 845 GACCGCGCTGGGAGGAGTACGCTGCAAGATTAACTCAAAAGAAATTGACGCGGAGAC 904  
 486 TCCGCGCTGGGAGGAGTACGCTGCAAGATTAACTCAAAAGAAATTGACGCGGAGAC 427  
 905 AAGCGGTGATATGTGATTAATTCGATGCAAGCGGAAACCTTAACCTACCTCTTGACA 964  
 426 AAGCGGTGATATGTGATTAATTCGATGCAAGCGGAAACCTTAACCTACCTCTTGACA 367  
 965 TGTAGCGAATTTTCTAGAGATAGATTAGT- CTTGGGAGCGCTTAACAGAGTGTGCTAT 1023  
 366 TGCGCGAATCCTCTTGAAGAAGAGGGGTGCTTCGCGGAGCGGAGACAGAGTGTGCTAT 307  
 1024 GCGTGTCTCACTGCTGTCTGAGATTTTGGGTTAACTCCCGGACGAGGCAACCCCTT 1083  
 306 GCGTGTCTCACTGCTGTCTGAGATTTTGGGTTAACTCCCGGACGAGGCAACCCCTT 247

Qy 1084 GTCATTAATTCGCATCATTTGGTGGGACCTTAAATGAGACTGCCGTGACAAACCGGAG 1143  
 Db 246 GTGTTTAGTTGCCACCGTTGAGTTTGAACCTGAAACACACTGCCGTGATTAAGCCGGAG 187  
 Qy 1144 GAAAGTGGGAGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGCTTACACAGTAATAC 1203  
 Db 186 GAAAGTGGGAGATGACGTCAAGTCTCATGAGCCCTTAAATGCGCTGGGAGACACAGTCTAC 127  
 Qy 1204 AATGCGCGCTACAGAGGCTTGGCAACCGCGAGGGGAGACTAATCTCAGAAACGGGCTG 1263  
 Db 126 AATGCGCGGAGACAAAGGTCGATCCCGAGGGTGAAGCTAATCTCAGAAACCGGCTCT 67  
 Qy 1264 TAGTCCGAGATCGGAGTCTGCAACTCGACTCCGTGAAGTGGAAATCGCTAATTCGCGGA 1323  
 Db 66 CAGTTTGAATTCAGAGGCTGCAACTCCCTGATGAAGCCGGAATCTCTAGTATCCGCCG 7  
 Qy 1324 TCAGC 1328  
 Db 6 TCAGC 2

RESULT 39  
 BH547523/c 875 bp DNA linear GSS 14-DEC-2001  
 LOCUS BOGH290TF BOGH Brassica oleracea genomic clone BOGH290, genomic  
 DEFINITION  
 SURVEY sequence.  
 ACCESSION BH547523 GI:17799304  
 VERSION BH547523.1  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 875)  
 Utterback,T.R., Mortman,J.R., White,O.R., and Town,C.D.  
 TITLE Whole genome shotgun sequencing of *Brassica oleracea* and its  
 application to gene discovery and annotation in Arabidopsis  
 JOURNAL Genome Res. 15 (4), 487-495 (2005)  
 PUBMED 15805490  
 COMMENT Other\_GSSs: BOGH290TR  
 Contact: Chris Town  
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TP  
 Class: sheared ends.

FEATURES  
 source  
 1. 875  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO100DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOGH290"  
 /clone\_1b="BOGH"  
 /note="Vector: pHOS1; Site\_1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN  
 Query Match 36.1%; Score 526.2; DB 9; Length 875;  
 Best Local Similarity 77.7%; Pred. No. 9e-151;  
 Matches 697; Conservative 0; Mismatches 173; Indels 27; Gaps 4;

Qy 235 GCTTACCAAGCGCATCATGATGTTGTTGAGAGACACCACTGGGACTGA 294  
 Db 873 GCTTACCAAGCGCATCATGATGTTGTTGAGAGATGATGACCACTGGGACTGA 814  
 Qy 295 GACACGCGCCAGACTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 354

Db 813 GACAGCGCCCAAGACTCTCTACGGAGGAGCAGATGGGGAATTTTCCGCAATGGCGCAAGC 754  
 QY 355 CTGATCCAGCATGCGCGCTGAGTGAAGAGCCCTTGGGTTTAAAGCTTTTCAGTGC 414  
 Db 753 CTGACGAGCAATGCGCGCTGAG-GTAGAAGGCTTACGGGTCTCTGAATCTTTTCCGAG 695  
 QY 415 AGAAGAAAAGTTACGGTAAATTAATGCTGACCGATCGACAGAGAAAGACCG 474  
 Db 694 AGAAG------AGCAATGACGATCTGGGGAATTAAGCATGC 658  
 QY 475 GCTAATCATGCTGACAGACCGCGTAAATCTAGAGGTGCAGACGGTAAATCGAATTAAT 534  
 Db 657 GCTAATCTGTGCAAGACCGCGTAAATCAAGATGCAACGTTATCCGAATGAT 598  
 QY 535 GGGCGTAAAGGCTGCGAGCGCGCTTTGATGTCAGATGTAATCCCGGCTTAACCT 594  
 Db 597 GGGCGTAAAGCTGCTGAGGTGCTTTTAAATCCCGCGTCAATCCAGGCTCAACCC 538  
 QY 595 GGGAAATTCGTTGAAAATCAAGCTAGATGTGCGACAGAGGAGGTGGAATTCATGTC 654  
 Db 537 TGACACAGCGCGTGAATCAAGCTTAGTACGGTGAAGGAGGGAATTTCCGATG 478  
 QY 655 TAGACGTAAATGCTAGATATGGAAGACATCGATGCGAAGCGCTCTGGGT 714  
 Db 477 GAGCGGTAAATGCTAGATGAGATCGAAAGAACACACAGCGCAAGCATCTGCTGGCC 418  
 QY 715 AACACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGTACCTGTAATC 774  
 Db 417 GACACTGACACTGAGACGAAAGCTAGAGGAGCGAATGAGTAAATCCCAAGTATGTC 358  
 QY 775 CAGCCCTAAAGATGCACTAGTGTGGGCTTATTAAGGCTTG--GTAAAGCAAGT 832  
 Db 357 CTAGCGGTAAACATGATATCTAGCGCTGTGCTATCGACCCGTGACGCTGATGCTA 298  
 QY 833 ACAGCTGAAGTATACCGCTGGGAGTACCGTGCAGAAATTAAGTCAAGAGAAATGAC 892  
 Db 297 ACCCGTTAAGTATCCCGCTGGGAGTACCTTGCAGAAATGAATCTCAAGAAATGAC 238  
 QY 893 GGGGACCCGCAAGCGGTGATTAATGATTAATTCGATGCAACGCGAAATCTTAC 952  
 Db 237 GGGGACCCGCAAGCGGTGATGATGATGATTAATTCGATGCAAGCGAAATCTTAC 178  
 QY 953 CTACCTTGAATGATGCAATTTCTAGATGATGATGATG--CTTCGGGAGCAAGTAA 1011  
 Db 177 CAGGCTTGAATGATGCGCAATCTCTTGAAGAGAGGCTCTTCGGGAGCGGAC 118  
 QY 1012 CAGGTCTGATGCTGTCTGACGCTGCTGCTGATGATGTTGGTTAAGTCCCGAAGC 1071  
 Db 117 CAGGTGCTGATGCTGTCTGACGCTGCTGCTGATGATGTTGGTTAAGTCCCGAAGC 58  
 QY 1072 AGCGCAACCTTGTCAATTAATGCTGATTTGGTGGGCACTTAAATGACCTGCC 1128  
 Db 57 AGCGCAACCTTGTCAATTAATGCTGATTTGGTGGGCACTTAAATGACCTGCC 1  
 RESULT 40  
 B2426275/c 889 bp DNA linear GSS 13-DEC-2002  
 LOCUS BOOAB81R BO.1.6.2 KB.101 Brassica oleracea genomic clone BOOAB81,  
 DEFINITION genomic survey sequence.  
 ACCESSION B2426275  
 VERSION B2426275.1 GI:26667787  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 Rosids; eustosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 889)  
 Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
 Uteback,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis

JOURNAL Genome Res. 15 (4), 487-495 (2005)  
 PUBMED 15805490  
 COMMENT Other\_GSSs: BOOAB81R  
 Contact: Chris Town  
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: 1R  
 Class: sheared ends.  
 FEATURES  
 source location/Qualifiers  
 1. 889  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000H3"  
 /db\_xref="taxon:3712"  
 /clone\_id="BOOAB81"  
 /note="Vector: pHD51; Site: 1.6-2 kb sheared total DNA inserted into pHD51 using BclXI linkers"  
 ORIGIN  
 Query Match 36.1%; Score 526; DB 9; Length 889;  
 Best Local Similarity 77.6%; Pred. No. 1e-150;  
 Matches 698; Conservative 0; Mismatches 175; Indels 27; Gaps 4;  
 QY 186 TCGCTTTTGAAGCGCGCGATGCTGATTAAGTCTGATGTTGGTGAAGGCTTAAAGCCTTAAAG 245  
 Db 876 TCGCGCCGAGAGGAGGCGCTGCGCTGATTAAGCTAG-TGGTAGCAAAATAGCTTAAACCAAG 818  
 QY 246 CCAAGATCATGATGTTGCTGTGAGAGACGACCAAGCAATCGGACTAGACACGCGCCA 305  
 Db 817 CCATATATCATGATGCTGCTCCAGAGATGATGACACATGGAATGAGACACGCGCCA 758  
 QY 306 GACTCTTACGAGGAGCAGACGATGGGAAATTTTGAACAATGGGCGCAAGCTGATCAGCA 365  
 Db 757 GACTCTTACGAGGAGCAGACGATGGGAAATTTTCCGCAATGGGCGAAACCTGACGAGCA 698  
 QY 366 ATGCGCGTGAATGAAGAGGCTTGGGTTGAAGCTTTTCACTGATGAGAAAGG 425  
 Db 697 ATGCGCGTGAATGAAGAGGCTTGGGTTGAAGCTTTTCCAGAGAAAG------ 643  
 QY 426 TTAAGTAAATATGCTGACCCATGACGATATGACAGAAAGACAGCGCTAACTAGT 485  
 Db 642 -----AGCAATGACGATATCTGGGAAATTAAGATGAGCTGACTCTGT 601  
 QY 486 GCGAGAGCGCGGATATCTGATGAGGTGCAAGCTTAAATCGGAATTAATGAGCGTAAAG 545  
 Db 600 GCGAGAGCGCGGATATCTGATGAGGTGCAAGCTTAAATCGGAATTAATGAGCGTAAAG 541  
 QY 546 GTGCGCAGCGCGCTTGAATGATGATGATGAATCCCGGCTTAACTGGGAATTCGT 605  
 Db 540 GTCTGATGCTGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481  
 QY 606 TTTGAATCTAAGAGTGTGCGACAGAGGAGGTGAAATTCATGCTGTAGCATGAA 665  
 Db 480 TTTGAATCTAAGAGTGTGCGACAGAGGAGGTGAAATTCATGCTGTAGCATGAA 421  
 QY 666 TCGGTATGATATGGAAGACATGATGCGAAGGAGCGCTCTGGGTTAACTGACGCG 725  
 Db 420 TCGGTATGATATGGAAGACATGATGCGAAGGAGCGCTCTGGGTTAACTGACGCG 361  
 QY 726 TCATCAGCAAGAGCGTGGGAGCAAAAGATTAATGATATCCCTGATGATGACCGCTTAA 785  
 Db 360 TGAAGAGCAAGAGCGTGGGAGCAAAAGATTAATGATATCCCTGATGATGACCGCTTAA 301  
 QY 786 CGATGCAATGATGTTGTTGGCTTTAATGATGATGATGATGATGATGATGATGATGATGAT 843  
 Db 300 CGATGATATCTGAGGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 241  
 QY 844 TGAAGCGCTGGGAGATGAGGTGCAAGATTAATGATCAAGAGAAATGAGCGGAGCCGCA 903

Db 240 ATCCCGCTGGGAGTAAGTTCGCAAGAACTCAAGATGACGCGGCGCGCA 181  
QY 904 CAGCGGTGATTAATGATTAATTCGATGCAACGGAAACCTTACCTTACCTTGAC 963  
Db 180 CAGCGGTGAGCATGTGTTTAATTCGATGCAAGGAAACCTTACCAAGGCTTGAC 121  
QY 964 ATGTAGCGAATTTTCTAGAGATGATTAAGTG-CTTCGGGAACGCTAACACAGGTGCTGCA 1022  
Db 120 ATCCCGGATTCCTCTTGAAAGAGAGGGGTGCTTCGGGAACGCGGACACAGGTGCTGCA 61  
QY 1023 TGGCTGTCTGACGCTGCTGCTGCAAGTGTGGGTTAAGTCCCGCAACGAGCGCAACCT 1082  
Db 60 TGGCTGTCTGACGCTGCTGCTGCAAGTGTGGGTTAAGTCAAGCAAGCGCAACCT 1

RESULT 41  
BH705306/c 873 bp DNA linear GSS 20-FEB-2002  
LOCUS BOMIV08TF BO 2.3 KB Brassica oleracea genomic clone BOMIV08,  
DEFINITION genomic survey sequence.  
ACCESSION BH705306  
VERSION BH705306.1 GI:18787817  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
REFERENCE Ayala, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
AUTHORS Utecherack, T.R., Mortman, J.R., White, O.R. and Town, C.D.  
TITLE Whole genome shotgun sequencing of Brassica oleracea and its  
JOURNAL application to gene discovery and annotation in Arabidopsis  
PUBMED Genome Res. 15 (4), 487-495 (2005)  
COMMENT Other GSSs: BOMIV08TR  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.  
FEATURES  
Location/Qualifiers  
Source 1..873  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="T0100DH3"  
/db\_xref="taxon:3712"  
/clone="BOMIV08"  
/clone\_id="BO\_2\_3\_KB"  
/note="Vector: pHOSt1, Site 1: BstXI, 2-3 kb sheared  
genomic DNA inserted into pHOSt1 using BstXI linkers"

Query Match 36.1%, Score 525.6, DB 9, Length 873;  
Best Local Similarity 77.5%, Pred. No. 1.4e-150;  
Matches 662, Conservative 0, Mismatches 189, Indels 3, Gaps 2;

QY 444 ACCCATGACGGTATCGACAGAAAGACCGGCTTACCTACGTCAGACGCGCGTAT 503  
Db 862 AGCAATGACGTATCTGGGAAATAGCATCGGCTACTCTGTGCGACAGCCGCGTAT 803  
QY 504 ACCTAGGAGTGAAGCGTTATCGAATTAATCTGGGCGTAAAGGGTCGCGAGCGGCTTGT 563  
Db 802 ACAGAGATGCAAGGCTTATCGGAATGATTGGGCGTAAAGCGTCTGATGAGTGGCTTTT 743  
QY 564 AAGTCAGATGGAATCCCGGCTTAACTGGGAAATGGCTTTGAAACTCAAGCTAG 623  
Db 742 AAGTCGCGCTCAATCCAGGGCTCAACCTTGACAGCGGCTGAAACTCAAGCTTG 683

QY 624 AGTGTGCAGAGGAGAGTGAATTCATGTTAGCAGTAAATGCTAGAGATAGGAAG 683  
Db 682 AGTACGATGGGGGACAGAGGAATTTCCGTTGAGCGCTTAATGCTTAAGATCGAAG 623  
QY 684 AACATGATGGCGAAGGACGCTCTGGGTTAACTGACGCTCATGACGAAAGCGTGG 743  
Db 622 AACACCAAGCGGGAAGCACTGCTGGGCCGACACTGACACTGACAGAGAAAGCTAG 563  
QY 744 GGAGCAACAGATTAATGATACCTGTTAGTCCAGCGCTTAACAGATGCTCAATGTTGT 803  
Db 562 GGAGCGAATGGGATTAATGATACCCAGTATGCTTACCGCTTAACAGATGATACTAGCGCT 503  
QY 804 GGGCTTATTAAGCTTG--GTAAAGAGCTAACGCGTGAAGTTGACCGCTGGGAGTAC 861  
Db 502 GTGCGTATGCAACCGGACAGTGTGATGCTAACGCGCTTAATGATTCGCGCTGGGAGTAC 443  
QY 862 GGTGCAAGATTAATCACTCAAGGAATGACGCGGACCCGACAGCGGTGATTAATGTG 921  
Db 442 GTTCGCAAGATTAATCACTCAAGGAATGACGCGGCGCCGACAGCGGTGATGATGTG 383  
QY 922 GATTAATTCGATGCAACGCGAAACCTTACTTACCTTGACATGTAGCGAATTTTCTAG 981  
Db 382 GTTAATTCGATGCAAGCGAAGCACTTACAGGCGCTTGACATGCGCGAATCTCTTG 323  
QY 982 AGATGATTAATG-CTTCGGGAACGCTAACACAGGTGCTGATGGCTGTCAGTCTGT 1040  
Db 322 AAAGAGAGGGGTCTTTCGGGAACGCGACACAGGTGTCATGAGCTGTGACGCTGT 263  
QY 1041 GTGCTGATGTTGGGTTAAGTCCCGCAACGACGCAACCTTGTCTTAATTTGCCATCA 1100  
Db 262 GCGGTAAAGTGTGGTTAAGTCCCGCAACGACGCAACCTTGTGTTAGTTGCCACG 203  
QY 1101 TTTGTTGGGCACTTTAATGAGACTGCGGTGCAAAACGAGAGAGTGGGATGACGT 1160  
Db 202 TTGAGTTTGAACCTTGAAACAGACTGCGGTGATTAAGCCGAGAGAGTGAAGTACGT 143  
QY 1161 CAAGTCTCATGCGCCCTTAATGAGTGGGCTTCAACGTAATCAATGCGCGTCAAGG 1220  
Db 142 CAAGTCAATCATGCGCCCTTAATGAGTGGGCTTCAACGTAATGAGGCGGACAAAG 83  
QY 1221 GTTGCACACCGGAGGAGGAGCTAATCTGAGAAAGCGCGTGTAGTCCGATGCAAGTC 1280  
Db 82 GTGCGATTCGCGGAGGAGTGAAGTCACTCAAAAACCGCTCAGTTCCGATTGCAAGC 23  
QY 1281 TGCAACTGCACTCC 1294  
Db 22 TGCAACTGCGCTGC 9

RESULT 42  
CF566443 745 bp mRNA linear EST 23-SEP-2003  
LOCUS CF566443  
DEFINITION 115107G11.x1 C. reinhardtii CG-1690 (mt+), Stress III  
(normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 3', mRNA  
sequence.  
ACCESSION CF566443  
VERSION CF566443  
KEYWORDS EST.  
SOURCE Chlamydomonas reinhardtii  
ORGANISM Chlamydomonas reinhardtii  
REFERENCE Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C.,  
AUTHORS Merchant, S., Quinn, J., and Shrager, J.  
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
JOURNAL Unicellular System for Analyzing Gene Function and Regulation in  
COMMENT unpublished (2003)  
Contact: Charles Hauser  
DCMB Box 91000  
Duke University



Durham, NC 27708-1000  
 Tel: 919 613 8159  
 Fax: 919 613 8177  
 Email: chauser@duke.edu  
 High quality sequence scop: 745.  
 Location/Qualifiers

## FEATURES

source

1..745  
 /organism="Chlamydomonas reinhardtii"  
 /mol\_type="mRNA"  
 /strain="CC-1690 wild type mt+ 21gr"  
 /db\_xref="taxon:3055"  
 /clone\_1lb="C. reinhardtii CC-1690 (mt+), Stress III  
 (normalized), lambda zap II"  
 /note="Vector: pBluescript II SK-; Site\_1: EcoRI, Site\_2:  
 XhoI; Stress condition III library, constructed by  
 Jeanette Quinn and Chung-wen Chang, combined cDNAs from  
 CC-1690 cells grown to mid-log phase in copper-free TAP  
 medium (see Quinn and Merchant (1998) Methods in  
 Enzymology, 299:263-279) in a shaking (250 rpm) illuminated  
 (about 100 micromole/m2/sec) incubator at 22 C (see Quinn  
 and Merchant (1998) Methods in Enzymology, 299:263-279);  
 CC-1690 cells grown to mid-log phase in low Fe (1  
 micromolar Fe) TAP medium (see La Fontaine S, Quinn JM,  
 Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J,  
 Merchant S, LaFontaine et al. (2002) Eukaryotic Cell,  
 1:736-757) in a shaking illuminated incubator (same  
 conditions as above). CC-1690 cells were grown to mid-log  
 phase in TAP medium in a shaking illuminated incubator to  
 a density of 8x10<sup>6</sup> cells/ml. The flask was transferred to  
 a shaking platform (200 rpm) at room temperature (23C) 12  
 micromole/m2/sec illumination and bubbled in a scopered  
 flask with 98% nitrogen, 2% CO<sub>2</sub> gas mixture for 1 hour  
 prior to harvesting for RNA isolation (as per Quinn JM,  
 Barraco P, Eriksson M, Merchant S, Quinn et al. (2000) JBC  
 275:6080-6089); CC-1690 cells grown to mid-log phase  
 (3x10<sup>6</sup> cells/ml) in TAP medium in a shaking (150 rpm)  
 illuminated (70 mole photon/m2/sec) incubator at 27 C.  
 Cells were diluted to 1x10<sup>6</sup> cells/ml, transferred to high  
 light (11000 mol photon/m2/s) with shaking (150 rpm) and  
 sampled at (0.5, 1.2, 4, 6, 12 hrs); PolyA mRNA was purified  
 from each sample, pooled and cDNA synthesized (see Strager  
 et al. Plant Physiol. 131, 401-408 for details). The cDNA  
 was directionally cloned into lambda zap II (Stratagene)  
 in the EcoRI (5') and XhoI (3') sites. pBluescript II SK-  
 plasmids were excised from the lambda Zap clones by  
 superinfection with Exsist (Stratagene) phage. The  
 library was normalized using method 4 described in Bonaldo  
 et al., (1996) Genome Research 6: 791-806."

## ORIGIN

Query Match 35.8%; Score 521.8; DB 6; Length 745;  
 Best Local Similarity 84.3%; Pred. No. 2e-149;  
 Matches 622; Conservative 0; Mismatches 113; Indels 3; Gaps 3;

468 AGACACGGCTAACTAGTCGACAGCCGGGTAATACGAGGAGGTTAATCGG 527  
 |||||  
 Db 8 AGACACGGCTAACTAGTCGACAGCCGGGTAATACGAGGAGGTTAATCGG 67  
 528 AATTACTGGGCGTAAAGGTCGACAGCCGGCTTGTAACTCAGATGTGAATCCCGGGC 587  
 |||||  
 Db 68 AATTACTGGGCGTAAAGGTCGACAGCCGGCTTGTAACTCAGATGTGAATCCCGGGC 127  
 588 TTAACCTGGGAATTCGCTTTGAAACTACAAAGCTAGAGTGTGCAAGAGGAGGTGAATT 647  
 |||||  
 Db 128 TCAACCTGGGAATTCGCTTTGAAACTACAAAGCTAGAGTGTGCAAGAGGAGGTGAATT 187  
 648 CCAATGTGTAGCACTGAATTCGCTTTGAAACTACAAAGCTAGAGTGTGCAAGAGGAGGTG 707  
 |||||

Db 188 CCAAGGTAGACGGGTGAATAATGCGTAGAGATCTGGAAGAAATACCGGTGCGAAGCGGCCCC 247  
 |||||  
 Qy 708 CTGGGTAACTAGACCTGATGCAAGAAAGCTGGGAGCAAAAGGATTAAGATACCT 767  
 |||||  
 Db 248 CTGACGAAAGACTGACGCTCAGATGCGAAAGGTGGGAGCAAAAGGATTAAGATACCT 307  
 |||||  
 Qy 768 GGTAGTCCACGCGCTTAAAGCATGATGCAATGATGTTGGGCTTATTAAGCTTGG-TAAAG 826  
 |||||  
 Db 308 GGTATGCAACGCGCTTAAAGCATGATGCAATGATGTTGGGCTTATTAAGCTTGG-TAAAG 367  
 |||||  
 Qy 827 AAGCTAACGCGTGAAGTTGACCGCTGGGAGTACGCTGCGCAAGATTAATTAATCAAGGA 886  
 |||||  
 Db 368 GAGCTAACGCGTGAAGTTGACCGCTGGGAGTACGCTGCGCAAGATTAATTAATCAAGGA 427  
 |||||  
 Qy 887 ATTGACGGGAGACCCGCAAGAGGCTGATTAATGATTAATGATGATGATGATGATGATGAT 946  
 |||||  
 Db 428 ATTGACGGGAGACCCGCAAGAGGCTGATTAATGATTAATGATGATGATGATGATGATGAT 487  
 |||||  
 Qy 947 CCTTACCTACCTTACATGTAAGCAATTTTCTAGAGAT-AGATTAGTCTTCGGGAAACG 1005  
 |||||  
 Db 488 CTTTACCTGCTTGTATGATTCACGGAAGTTTCAAGATGAGATGTGCTTCGGGAAAC 547  
 |||||  
 Qy 1006 CTACACAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065  
 |||||  
 Db 548 GTGAGACAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607  
 |||||  
 Qy 1066 GCAAGAGGCGCAACCTTGTATTAATGCTATC-ATTGCTTGGGCACTTAAATGAGAC 1124  
 |||||  
 Db 608 GCAAGAGGCGCAACCTTGTATTAATGCTATC-ATTGCTTGGGCACTTAAATGAGAC 667  
 |||||  
 Qy 1125 TSCCGGTGCAAAACCGAGAGAGTGGGATGACGTCAGTCCCTTATGAGT 1184  
 |||||  
 Db 668 TSCCGGTGCAAAACCGAGAGAGTGGGATGACGTCAGTCCCTTATGAGT 727  
 |||||  
 Qy 1185 AGGCGTTACACGTAATA 1202  
 |||||  
 Db 728 AGGCGTTACACGTAATA 745  
 |||||

RESULT 43  
 BH482629 842 bp DNA linear GSS 13-DEC-2001  
 LOCUS BGNK19TR BGN Brassica oleracea genomic clone BGNK19, genomic  
 DEFINITION survey sequence.  
 ACCESSION BH482629.1 GI:17690733  
 VERSION BH482629.1  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## PUBMED

## COMMENT

15805490  
 Other GSSs: BGNK19TF  
 Contact: Chris Town  
 TIGR  
 5712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.

## FEATURES

source

1..842  
 Location/Qualifiers  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"



## ORIGIN

/strain="T01000DH3"  
/db\_xref="taxon:3712"  
/clone="BOGK19"  
/clone\_1lb="BOGK"  
/note="Vector: pHO5; Site 1: BstXI; 2-3 kb sheared  
genomic DNA inserted into pHO5 using BstXI linkers"

35.8%; Score 521.6; DB 9; Length 842;

Query Match Best Local Similarity 77.8%; Pred. No. 2,4e-149; Matches 655; Conservative 0; Mismatches 184; Indels 3; Gaps 2;

469 GCACCGGCTAACTACGTGCGACGACCGCGGTAACTAGTGGTCAAGCGTTAATCGA 528  
1 GCATCGGCTAACTGTGTGCGACGCGCGGTAACTACAGAGATGCAAGCGTTATCCGGA 60  
529 ATTAAGTGGGCTAAAGGGTGGCGCGGCTTGTAGTCAAGTGTAAATCCCGGGCT 588  
61 ATGATTTGGGCGTAAAGCGTGTGTAGTGGCTTTTAAAGTCCGCGTCAAAATCCAGGGCT 120  
589 TAACCTGGGAATTGCGTTTGAATACAAAGCTAGATGTGGCAGAGGGAGTGAATTC 648  
121 CAACCTTGAGACAGCGGTGGAATCTACCAAGCTTGAATCGTATGGGGCAGAGGAAATT 180  
649 CATGTGTAGCAGTGAATGCGTAGAGATATGAAGAAATCATGATGGCGAAGCGACTCC 708  
181 CCGGTGAGCGGTGAATGCGTAGAGATATGAAGAAATCAACACCGCGAAGCACTCTGC 240  
709 TGGGTAACTAGACGCTCATGACAGAAAGCGTGGAGCAAAAGATTAATGATCCTG 768  
241 TGGGCCCACTGACACTGACAGAGAAAGCTAGAGGGGCAAAATGGGATTAATGATACCCA 300  
769 GTAGTCCACGCGCTTAACAGATGTCACTAGTTGTTGGGCTTATTAAGGCTTG--GTAAAG 826  
301 GTAGTCCACGCGCTTAACAGATGTCACTAGTTGTTGGGCTTATTAAGGCTTG--GTAAAG 826  
827 AAGCTAAACGCTGAAGTTGACCGCTGGGAGTACCGGTGCAAGATTAATCAAGGA 886  
361 TAGCTAAGCGTTAAGTATCCCGCTGGGAGTACGTTCCGAAAGATTAATCAAGGA 420  
887 ATTGACGGGAGCCGCGACAGCGGTGATTAATGTGATTAATTTGATGCAACGGGAAAA 946  
421 ATTGACGGGAGCCGCGACAGCGGTGATTAATGTGATTAATTTGATGCAACGGGAAAA 946  
947 CCTTACCTACCTTGAATGATGAGGAATTTCTAGAGATAGATTAAGT--CTTCCGGAAG 1005  
481 CCTTACCTACCTTGAATGATGAGGAATTTCTAGAGATAGATTAAGT--CTTCCGGAAG 1005  
1006 CTTAACACAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065  
541 CGGACACAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
1066 GCAAAGAGCGCAACCTTGTGATTAATGATGCACTTTGTTGGGCACTTAATGAGCT 1125  
601 GCAAAGAGCGCAACCTTGTGATTAATGATGCACTTTGTTGGGCACTTAATGAGCT 1125  
1126 GCGGTGCAAAACCGGAGAGAGTGGGATGACGTCATGCTTCAATGGCCCTTATGGGTA 1185  
661 GCGGTGCAAAACCGGAGAGAGTGGGATGACGTCATGCTTCAATGGCCCTTATGGGTA 1185  
1186 GGGCTTACACAGTAACTAATGATGCGCTGACAGAGGGTGGCCAAACCGGAGGGGAGCTA 1245  
721 GGGGCAACACAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
1246 ATCTGAGAAAGCGCGTGTAGTCCGATGCGAGTGTGCACTGTGACTCCGTGAAGTCCGA 1305  
781 ACTCAAAAGAGCGCGTGTAGTCCGATGCGAGTGTGCACTGTGACTCCGTGAAGTCCGA 1305  
1306 AT 1307  
841 AT 842

## RESULT 44

B2483676

LOCUS

BOOAS16TR\_BO\_1.6\_2\_KB\_tot

DEFINITION

Brassica oleracea genomic clone BOOAS16,

genomic survey sequence.

ACCESSION

B2483676

VERSION

B2483676.1 GI:26786074

KEYWORDS

SSS.

SOURCE

Brassica oleracea

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (base 1 to 837)

Ayele, M., Haas, B.D., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,

Utecherback, T.R., Wortman, J.R., White, O.R. and Town, C.D.

Whole genome shotgun sequencing of Brassica oleracea and its

application to gene discovery and annotation in Arabidopsis

Genome Res. 15 (4), 487-495 (2005)

15805490

Other SSSs: BOOAS16TF

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES  
source  
Location/Qualifiers  
1..837  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="T01000DH3"  
/db\_xref="taxon:3712"  
/clone\_1lb="BO\_1.6\_2\_KB\_tot"  
/note="Vector: pHO5; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pHO5 using BstXI linkers"

## ORIGIN

Query Match 35.8%; Score 521.4; DB 9; Length 837;

Best Local Similarity 78.0%; Pred. No. 2,7e-149; Matches 653; Conservative 0; Mismatches 181; Indels 3; Gaps 2;

469 GCACCGGCTAACTACGTGCGACGACCGCGGTAACTAGTGGTCAAGCGTTAATCGA 528  
1 GCATCGGCTAACTGTGTGCGACGCGCGGTAACTACAGAGATGCAAGCGTTATCCGGA 60  
529 ATTAAGTGGGCTAAAGGGTGGCGCGGCTTGTAGTCAAGTGTGAATCCCGGGCT 588  
61 ATGATTTGGGCGTAAAGCGTGTGTAGTGGCTTTTAAAGTCCGCGTCAAAATCCAGGGCT 120  
589 TAACCTGGGAATTGCGTTTGAATTAACAAAGCTAGATGTGGCAGAGGGAGTGAATTC 648  
121 CAACCTTGAGACAGCGGTGGAATTAACAAAGCTAGATGTGGCAGAGGGAGTGAATTC 180  
649 CATGTGTAGCAGTGAATGCGTAGAGATATGAAGAAATCATGATGGCGAAGCGACTCC 708  
181 CCGGTGAGCGGTGAATGCGTAGAGATATGAAGAAATCAACACCGGAGAAAGCACTCTGC 240  
709 TGGGTAACTAGACGCTCATGACAGAAAGCGTGGAGCAAAAGATTAATGATCCTG 768  
241 TGGGCGACACTGACACTGACAGAAAGCTAGAGGGAGCAAAATGGGATTAATGATACCCA 300  
769 GTAGTCCACGCGCTTAACAGATGTCACTAGTGTGGGCTTATTAAGGCTTG--GTAAAG 826  
301 GTAGTCCACGCGCTTAACAGATGTCACTAGTGTGGGCTTATTAAGGCTTG--GTAAAG 826  
827 AAGCTAAACGCTGAAGTTGACCGCTGGGAGTACCGGTGCAAGATTAATCAAGGA 886  
361 TAGCTAAGCGTTAAGTATCCCGCTGGGAGTACGTTCCGAAAGATTAATCAAGGA 420



SOURCE Chlamydomonas reinhardtii  
 ORGANISM Chlamydomonas reinhardtii  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;  
 Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.  
 1 (bases 1 to 770)  
 REFERENCE Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C.,  
 Merchant, S., Quinn, J., and Strager, U.  
 Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
 Unicellular System for Analyzing Gene Function and Regulation in  
 Vascular Plants. Project: 1115  
 Unpublished (2003)  
 JOURNAL Contact: Charlee Hauser  
 COMMENT DMB Box 91000  
 Duke University  
 Durham, NC 27708-1000  
 Tel: 919 613 8159  
 Fax: 919 613 8177  
 Email: chauser@duke.edu  
 High quality sequence stop: 770.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..770  
 /organism="Chlamydomonas reinhardtii"  
 /mol\_type="mRNA"  
 /strain="CC-1690 wild type mt+ 21gr"  
 /db\_xref="taxon:3055"  
 /clone\_id="C. reinhardtii CC-1690 (mt+), Stress III  
 (normalized), Lambda Zap II"  
 /note="Vector: pBluescript II SK-. Site 1: EcoRI; Site 2:  
 XhoI; Stress condition III library, combines cDNAs from  
 Jeanette Quinn and Chung-Wen Chang, combines cDNAs from  
 CC-1690 cells grown to mid-log phase in copper-free TAP  
 medium (see Quinn and Merchant (1998) Methods in  
 Enzymology, 2997:263-279) in a shaking (250 rpm) illuminated  
 (about 100 micromole/m2/sec) incubator at 22 C (see Quinn  
 and Merchant (1998) Methods in Enzymology, 2997:263-279);  
 CC-1690 cells grown to mid-log phase in low Fe (1  
 micromolar Fe) TAP medium (see La Fontaine S, Quinn JM,  
 Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J,  
 Merchant S, LaFontaine et al. (2002) Eukaryotic Cell,  
 1:736-757) in a shaking illuminated incubator (same  
 conditions as above). CC-1690 cells were grown to mid-log  
 phase in TAP medium in a shaking illuminated incubator to  
 a density of 8x10<sup>6</sup> cells/ml. The flask was transferred to  
 a shaking platform (200 rpm) at room temperature (23C) 12  
 micromole/m2/sec illumination and bubbled in a stoppered  
 flask with 98% nitrogen. 2% CO<sub>2</sub> gas mixture for 1 hour  
 prior to harvesting for RNA isolation (as per Quinn JM,  
 Barraco P, Eriksson M, Merchant S, Quinn et al. (2000) JBC  
 275:6080-6089); CC-1690 cells grown to mid-log phase  
 (3x10<sup>6</sup> cells/ml) in TAP medium in a shaking (150 rpm)  
 illuminated (70 mole photon/m2/sec) incubator at 27 C.  
 Cells were diluted to 1x10<sup>6</sup> cells/ml, transferred to high  
 light (11000 mol photon/m2/s) with shaking (150 rpm) and  
 sampled at (0.5, 1, 2, 4, 6, 12 hrs); CC-1690 cells grown to  
 mid-log phase in HS medium in a shaking (150 rpm)  
 illuminated (70 mole photon/m2/sec) incubator at 27 C.  
 Cells were diluted to 1x10<sup>6</sup> cells/ml, transferred to high  
 light (11000 mol photon/m2/s) with shaking (150 rpm) and  
 sampled at (0.5, 1, 2, 4, 6, 12 hrs). Poly A mRNA was purified  
 from each sample, pooled and cDNA synthesized (see Strager  
 et al, Plant Physiol. 131, 401-408 for details). The cDNA  
 was directionally cloned into lambda Zap II (Stratagene)  
 in the EcoRI (5') and XhoI (3') sites. pBluescript II SK-  
 plasmids were excised from the lambda Zap clones by  
 superinfection with Exsist (Stratagene) phage. The  
 library was normalized using method 4 described in Bonaldo  
 et al., (1996) Genome Research 6: 791-806."

Qy 468 AGCACCAGGCTAACTAGCTGCCAGACGCCGGTAATACCTAGGTCGCAAGCGTTAATCGG 527  
 |||||  
 Db 21 AACACCGGCTAACTCCGTCCAGACGCCCGGTAAATACGAGGGTCCAGCGTTATCCG 80  
 Qy 528 AATTACTGGGCGTAAAGGGTGGCGAGCGCGCCCTTGAAGTCAGATGTGAATCCCGGGC 587  
 |||||  
 Db 81 AATTACTGGGCGTAAAGGCCGACGACGCGGTTTGAATGAGATGTGAATCCCGGGC 140  
 Qy 588 TTAACCTGGGAATTCGCTTTGAACCTCAAAAGCTAAGTGTGCAAGGAGGTGAAT 647  
 |||||  
 Db 141 TCAACCTGGGAATTCGATCTGATCTGAGTCAAGCTTGAAGTCTGTAAAGGGGGTGAAT 200  
 Qy 648 CCATGTGTGACAGTGAATATGCGTAGATATGGAAGAAATCATGATGCGGACGACCTC 707  
 |||||  
 Db 201 CCAAGTGTAGCGGTGAATATGCGTAGATCTGAGGAATACCGGTGCGAAGCGGCGCC 260  
 Qy 708 CTGGGTTAACTAGTACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCT 767  
 |||||  
 Db 261 CTGGAGGAAGATGACGCTCAGGTGCAAAAGCGTGGGAGCAAAACAGATTAGATACCT 320  
 Qy 768 GGTAGTCCAGCGCTTAAAGATGTCAACTAGTGTGGGCTTATTAAGCTTGG-TAAG 826  
 |||||  
 Db 321 GGTAGTCCAGCGCTTAAAGATGTCAACTAGTGTGGGCTTATTAAGCTTGG-TAAG 380  
 Qy 827 AAGCTAACGGGTGAAGTGAACGCGCTGGGGAGTACGCTGCAAGATTAAACTCAAGGA 886  
 |||||  
 Db 381 GAGCTAACGGGTGAAGTGAACGCGCTGGGGAGTACGCGCTGCAAGATTAAACTCAAGGA 440  
 Qy 887 ATTGACGGGGAGCCCGACCAACCGGTGATTTATGATGATGCAACCGGAAAA 946  
 |||||  
 Db 441 ATTGACGGGGAGCCCGACCAACCGGTGATTTATGATGATGCAACCGGAAAA 500  
 Qy 947 CTTTACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1005  
 |||||  
 Db 501 CTTTACCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 560  
 Qy 1006 CTTAACAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1065  
 |||||  
 Db 561 GTGAGAGAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 620  
 Qy 1066 GCAACGAGCCCAACCCCTTGCATTAATTTGCATC-ATTGGTTGGGCACTTTAATGAGAC 1124  
 |||||  
 Db 621 GCAACGAGCCCAACCCCTTGCATTAATTTGCATC-ATTGGTTGGGCACTTTAATGAGAC 680  
 Qy 1125 TGGCGGTGACAAACCGGAGAGGTGGGATGACGTCAATCTCATGCGCTTTATGGGT 1184  
 |||||  
 Db 681 TGGCAGTGA-TAATGAGAGAGAGGTGGGATGACGTCAATCTCATGCGCTTTATGAGCC 739  
 Qy 1185 AGGCGTTCAACGTAATTAATGAGAGGTGGGATGACGTCAATCTCATGCGCTTTATGAGCC 1215  
 |||||  
 Db 740 AGGCGTTCAACGTAATTAATGAGAGGTGGGATGACGTCAATCTCATGCGCTTTATGAGCC 770  
 |||||  
 RESULT 47  
 BH700021/c 995 bp DNA linear GSS 20-FEB-2002  
 LOCUS BOMND50TR BO 2.3 KB Brassica oleracea genomic clone BOMND50,  
 DEFINITION genomic survey sequence.  
 ACCESSION BH700021  
 VERSION BH700021.1 GI:18775982  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 995)  
 REFERENCE Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
 Utecher, T.R., Wortman, J.R., White, O.R. and Yom, C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 JOURNAL PUBMED  
 15805490

## COMMENT

Other\_GSSs: BOMND50TF  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.

## FEATURES

source

Location/Qualifiers  
 1..995  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO100DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOMND50"  
 /clone\_1kb="BO.2.3 KB"  
 /note="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHO51 using BstXI linkers"

## ORIGIN

Query Match 35.7%; Score 520; DB 9; Length 995;  
 Best Local Similarity 74.8%; Pred. No. 7,7e-149;  
 Matches 732; Conservative 0; Mismatches 207; Indels 39; Gaps 5;

64 GTGGCAGATGGCGGACGGGTAGTATCATCGAAGCGTAAAGCGGGGTTAAGC 123  
 942 GTTTCAGATGGCGGACGGGTAGTATCATCGAAGCGTAAAGCGGGGTTAAGC 883  
 124 CATCGAAGATGTGCTAATACCCCATATATCTAAGAGAGAAAGCGGGATCGAAGAC 183  
 882 GCTGGAAGACGGCTGCTAATACCCCATATATCTAAGAGAGAAAGCGGGATCGAAGAC 832  
 184 CTGCGCTTTTGGAGGCGCGCATGTCTGATTAGCTAGTGGGGTAAAGGCTTACCA 243  
 831 ----CGCCGAGAGAGGGGCTCGGCTCTGATTACTGTTGTTAGGCAATAGCTTACCA 776  
 244 GGGCAGATCATAGTGTGCTGAGAGAGACGACCAAGCACTGGGACTGAGACAGCGCC 303  
 775 GGGGATGATCATAGTGTGCTGAGAGAGATCAAGCACTGGGACTGAGACAGCGCC 716  
 304 CAGACTCTTACGGGAGGCGCAATGGGGAATTTTGGACATAGGGCCGACCTGACAG 363  
 715 CAGACTCTTACGGGAGGCGCAATGGGGAATTTTGGACATAGGGCCGACCTGACAG 656  
 364 CAATGCGCGTGAAGTAAAGGCTTCGGGTTTAAAGCTCTTTCAGTGAAGAGAA 423  
 655 CAATGCGCGTGAAGTAAAGGCTTCGGGTTTAAAGCTCTTTCAGTGAAGAGAG 599  
 424 GGTTCAGTAAATATCTGAGACCCATGACGGTATCGACAGAGAGACCGGCTAACTAC 483  
 598 -----AGCAATGACGGTATCTGGGGAATAGAGATGGGCTTAACTCT 559  
 484 GTGCCAGACGCCCGGTAAATAGTAGGTGCAAGCGTAACTCGAATTAATGGGCGTAA 543  
 558 GTGCCAGACGCCCGGTAAATAGTAGGTGCAAGCGTAACTCGAATTAATGGGCGTAA 499  
 544 GGGTGGCAGAGCGGCTTGTAAATGATGTAATCCCGGCTTAACTCGGAAATTCG 603  
 498 GCGTCTGTAGGTGGCTTTTAAATGATGTAATCCCGGCTTAACTCGGAAATTCG 439  
 604 GTTGAAGACTCAAGCTTAGAGTGTGGCAGAGGGAGGTGAATTCATGTGTAGCAGTGA 663  
 438 GGTGAAGACTCAAGCTTAGAGTGTGGCAGAGGGAGGTGAATTCATGTGTAGCAGTGA 379  
 664 AATGCTGAAGATATGAGAGACATGATGGGAGGACCTCTCGGTTAACTCACTGAC 723  
 378 AATGCTGAAGATATGAGAGACATGATGGGAGGACCTCTCGGTTAACTCACTGAC 319  
 724 GCTCATGACGAAGAGCGTGGGAGCAACAGATTAATACCTGTGATGACAGCGCTTA 783  
 318 ACTGAGAGAGCAAGCTAGGGGAGCGAATGGGATTAATACCTCAATGATGCTAGCGCTTA 259

QY 784 AACGATGCACTAGTGTGGGCTTATTAAGGCTTG--GTAAAGAGTAAAGCGGTGA 841  
 DB 258 AACGATGCACTAGTGTGGGCTTATTAAGGCTTG--GTAAAGAGTAAAGCGGTGA 199  
 QY 842 GTTGAACCGCTGGGAGAGTACGGTCCGAAGATTAAATCAAGAAATTGACGGGACCCG 901  
 DB 198 GTATCCCGCTGGGAGAGTACGGTCCGAAGATTAAATCAAGAAATTGACGGGACCCG 139  
 QY 902 CACAAGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAACCTTACTACCTTG 961  
 DB 138 CACAAGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAACCTTACTACCTTG 79  
 QY 962 ACATGATGCAATTTTCTAGATGATTAATG--CTTGGGGAACGCTAACAGAGTCTG 1020  
 DB 78 ACATGATGCAATTTTCTAGATGATTAATG--CTTGGGGAACGCTAACAGAGTCTG 19  
 QY 1021 CATGGCTGTCTGACCTC 1038  
 DB 18 CATGGCTGTCTGACCTC 1

RESULT 48  
 BH685417 1020 bp DNA linear GSS 19-FEB-2002  
 LOCUS BOMFG90TR BO.2.3 KB Brassica oleracea genomic clone BOMFG90,  
 DEFINITION genomic survey sequence.  
 ACCESSION BH685417  
 VERSION BH685417.1 GI:18755854  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 1020)  
 Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
 Utecherback,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 15805490

JOURNAL  
 PUBMED  
 Other\_GSSs: BOMFG90TF  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.

## FEATURES

source

Location/Qualifiers  
 1..1020  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO100DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOMFG90"  
 /clone\_1kb="BO.2.3 KB"  
 /note="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHO51 using BstXI linkers"

## ORIGIN

Query Match 35.6%; Score 519.2; DB 9; Length 1020;  
 Best Local Similarity 73.9%; Pred. No. 1.4e-148;  
 Matches 743; Conservative 0; Mismatches 223; Indels 40; Gaps 5;

64 GTGGCAGATGGCGGACGGGTAGTATCATCGAAGCGTAAAGCGGGGTTAAGC 123  
 47 GTTTCAGATGGCGGACGGGTAGTATCATCGAAGCGTAAAGCGGGGTTAAGC 106  
 124 CATCGAAGATGTGCTAATACCCCATATATCTAAGAGAGAAAGCGGGATCGAAGAC 183

Db	107	GCTGGAAACGGCTCTTAATACCCCGTAG--CTAGAGACAAAGAGAAATC-----	157
Qy	184	CTTCCGCTTTTGGAGCGGCCGAGTGTGATTAGTAGTGGTGAAGGCTTACCA	243
Db	158	----CGCCGAGAGAGGGGCTCGCGTCTGATTAGTAGTGGTAGCAATAGCTTACAA	213
Qy	244	GCGCAGCATCATGTAGTTGGTCTTGAGAGACGACCAAGCCACACTGGGACTGAGACACGGCC	303
Db	214	GGCGATGATCATGTAGCTGGTCCGAGAGGATGATCAGCCACTGGGACTGAGACCGGCC	273
Qy	304	CAGACTCTTACGGGAGCGACGAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATCAG	363
Db	274	CAGACTCTTACGGGAGCGACGAGTGGGAAATTTTCCGAAATGGGCGAAAGCTTACCGAG	333
Qy	364	CAATGCCCGCGTAGTGAAGAGGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAAAGAA	423
Db	334	CAATGCCCGCGTAGTGAAGAGGCTTACGGTCTCTGAACCTTTTCCGAGAGAA--	390
Qy	424	GGTTACGGTAAATATCGTGAACCATGACGGTATCGACAGAGAGACCGGCTTAATCAG	483
Db	391	-----AGCAATGACGGTATCTGGGAGATAGCATCGGCTAATCTCT	430
Qy	484	GTGCGACAGACCGCGGTATATACGTAGGCTGCAAGGCTTAATCGGAATTACTGGGCGTAA	543
Db	431	GTGCGACAGACCGCGGTATATACAGAGATGCAACGCTTATCCGGAATGATTGGGCGTAA	490
Qy	544	GGGTGCGAGGCGGCTCTGTAAATGTCAGATGTGAATCCCGGGGCTTAACCTGGGAAATTGC	603
Db	491	GCGCTGTAGGTGGTGGTTTTTAAGTCCGCGTCAATCCAGGGCTCAACCTTGACAGGC	550
Qy	604	GTTTGAATCTCAAAAGCTAGAGTGTGGCAGAGGGAGTGAATTCATGTAGTACGTGA	663
Db	551	GGTGAATCTCAAGCTTGAAGTACGTAGGGGACAGAGGAATTTCCGGTGAAGCGTGA	610
Qy	664	AATGCGTAGAGATATGAAAGACATCGATGCGAAGGAGCGCTCTGGGTTAACACTGAC	723
Db	611	AATGCGTAGAGATCGAAGAAACACCAACGCGAAGAACCTCTGCTGGGCGGACACTGAC	670
Qy	724	GCTCATGCACGAAAGCGTGGGAGGAAACAGATTAAGTACCCTGTGTAGTCCAGGCCCTA	783
Db	671	ACTGAGAGACGAAAGCTAGGGGAGGAAATGGGATTAAGTATACCCGATGTCTTAGCCCTGA	730
Qy	784	AACGATGTCATCTAGTGTGGGCTTATTTAGGCTTG--GTAAACGAACTAACCGGTAA	841
Db	731	AACGATGATTACTAGGCGCTGTGCTATTCGACCCTGCAATGCTGTAGCTAACCGGTTAA	790
Qy	842	GTTGACCCGCTGGGGAGTAGCGGTCCGCAAGATTAAACTCAAGGAATTGACGGGACCG	901
Db	791	GTAATCCGCTGGGGAGTAGTCTGCCAAGAAATGAACTTCAGAGGAATTGACGGGGCCCG	850
Qy	902	CACAGACGGTGAATTATGTGATTAAATTCGATGCAACGCGAAGAACTTTACTACCTTGG	961
Db	851	CACAGACGGTGAAGCATGTGTTAATTGATGCAAGGAAGGAAGACTTTACAGAGGCTTG	910
Qy	962	ACATGTAGGAAT--TTTCTAGAGATAGATTAGTCTTCGGGAAGCTTAACAGAGTCT	1019
Db	911	ACATGCCCGGAATCCCTCTTGAAAGAGGGGTGCTTTCGGAAGCGCGACACAGTGGT	970
Qy	1020	GCATGGCTGTGCTCAGACTCGTGCATGAGATTGGGTTAAAGTCC	1065
Db	971	GCATGGCTGTGCTCAGCTGCTGCCGTAAGTGTGGGTTAAAGTCC	1016

RESULT	49
BH545929	
LOCUS	849 bp DNA linear GSS 14-DEC-2001
DEFINITION	BOGNL60TF BOGN Brassica oleracea genomic clone BOGNL60, genomic survey sequence.
ACCESSION	BH545929
VERSION	BH545929.1 GI:17797710
KEYWORDS	GSS.
SOURCE	Brassica oleracea

ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT
Brassica oleracea	1 (bases 1 to 849)			
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Brassicales; Brassicaceae; Brassica				
Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utecherack, T.R., Wortman, J.R., White, O.R. and Town, C.D.				
Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis				
Genome Res. 15 (4), 487-495 (2005)				
15805490				
Other GSSs: BGNL60TR				
Contact: Chris Town				
TIGR				
9712 Medical Center Drive, Rockville, MD 20850, USA.				
Tel: 301-838-3523				
Fax: 301-838-0208				
Email: cdtown@tigr.org				
DNA is from a doubled haploid provided by Tom Osborn.				
Seq primer: TP				
Class: sheared ends.				
FEATURES				
source	Location/Qualifiers			
	1..849			
	/organism="Brassica oleracea"			
	/mol_type="genomic DNA"			
	/strain="TO1000DH3"			
	/db_xref="taxon:3712"			
	/clone="BGNL60"			
	/clone_11b="BGN"			
	/note="Vector: pHOS1; site_1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"			
ORIGIN				
Query Match	35.6%;	Score 518.8;	DB 9;	Length 849;
Best local Similarity	78.1%;	Pred. No. 1.7e-148;		
Matches	662;	Conservative	0;	Mismatch 182; Indels 4; Gaps 3;
Db	516	AGCGTTATCGGAATTACTGAGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	575	
Qy	2	AGCGTTATCGGAATTACTGAGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	575	
Db	516	AAATCCCGGCGCTTAACTGAGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	635	
Qy	62	AAATCCCGGCGCTTAACTGAGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	635	
Db	62	AAATCCCGGCGCTTAACTGAGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	635	
Qy	636	GAGGTGGAATTCATGTGTAGCACTGAGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	695	
Db	122	GAGGTGGAATTCATGTGTAGCACTGAGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	695	
Qy	696	GAAGGACACCTCTGAGGATTAACTGAGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	755	
Db	182	GAAGGACACCTCTGAGGATTAACTGAGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	755	
Qy	756	ATTAGATACCTGTGATGTCACGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	815	
Db	242	ATTAGATACCTGTGATGTCACGCGCTTAAAGGATGCGAGCGCGCTTGTAGTCAATGTG	815	
Qy	816	GCTTG-CTAAGCAAGCTTAACGCGGATGTAACCGCTGAGGAGTACGTCGCAAGAT	873	
Db	302	CGTGCAATGTGTAAGCTTAACGCGGATGTAACCGCTGAGGAGTACGTCGCAAGAT	873	
Qy	874	AAATCTCAAAGGAATTGACGGGACCCGACAGCGGTGATTTATGTGATTAATTGAT	933	
Db	362	GAATCTCAAAGGAATTGACGGGACCCGACAGCGGTGATTTATGTGATTAATTGAT	933	
Qy	934	GCAAGCGAATACTTAACCTTACCTTGAATGAGCAATTTTCTAAGATAGTTAGT	993	
Db	422	GCAAGCGAATACTTAACCTTACCTTGAATGAGCAATTTTCTAAGATAGTTAGT	993	
Qy	994	G-CTTCGGGAACGCTTAACAGGATGTCATAGGCTGTGTAGCTCGTGTGTCGATGAT	1053	
Db	482	GCTTCGGGAACGCTTAACAGGATGTCATAGGCTGTGTAGCTCGTGTGTCGATGAT	1053	

QY 1053 TGGGTTAAGTCCCGCAACGAGCCGCTTGTCAATTAATGTCATCATTTGGTGGGA 1112  
 DB 542 TGGGTTAAGTCCCGCAACGAGCCGCTTGTCAATTAATGTCATCATTTGGTGGGA 601  
 QY 1113 CTTTAATGAGACTGCGGTGACAAACCGAGAGAGGTGGGATGATGATCAAGTCTCATG 1172  
 DB 602 CCTGAACAGACTGCGGTGATTAAGCGGAGAGAGATGATGATCAAGTCAAGTCAAG 661  
 QY 1173 GCCCTTATGAGTGGGCTTCAACGTAATCAATGCGCGTACAGAGGTTGCCAACCG 1232  
 DB 662 CCCCTTATGCGCGGTGACACAGTGTCTAATGCGCGGACAAAGGTCGCGATCCG 721  
 QY 1233 CAGAGGGGAGCTTAATCTCAGAAAGCGCGTGTAGTCCGATCGGATCGCAATGAGCT 1292  
 DB 722 CAGAGGGTGAAGCTTAATCTCAGAAACCGGTCTCAGTTCGATTCAGAGCTGCAATGCGCT 781  
 QY 1293 CCGTGAAGTCGAAATGCTAGTAATCGCGGATCAG-CATGTCGGGTGAAATGCTCCG 1351  
 DB 782 GCATGAAGCCCGATGCTAGTAATGCGCGGATCAGCATACGCGGTGAATTCGTCGCG 841  
 QY 1352 GGTCTTGT 1359  
 DB 842 GGCCTGT 849

RESULT 50  
 BH730827/c 848 bp DNA linear GSS 20-FEB-2002  
 LOCUS BOMGL56TF BO 2.3 KB Brassica oleracea genomic clone BOMGL56,  
 DEFINITION genomic survey sequence.  
 ACCESSION BH730827  
 VERSION BH730827.1 GI:18836222  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eustoids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 848)  
 Ayele,M., Haeß,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
 Utecherback,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 genome Res. 15 (4), 487-495 (2005)  
 JOURNAL PUBMED 15805490  
 COMMENT Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seg primer: TF  
 Class: sheared ends.

FEATURES  
 source  
 1..848  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOMGL56"  
 /clone\_lib="BO\_2.3 KB"  
 /note="Vector: pHOST1, Site 1: BclXI, 2-3 kb sheared  
 genomic DNA inserted into pHOST1 using BclXI linkers"

ORIGIN  
 Query Match 35.6%; Score 518.6; DB 9; Length 848;  
 Best Local Similarity 77.6%; Pred. No. 2e-148;  
 Matches 676; Conservative 0; Mismatches 169; Indels 26; Gaps 3;  
 QY 208 TCTGATTAGCTAGTGGTGGGTAAGGCTTACCAAGGCGAGCATGATGATGCTCTGA 267  
 DB 848 TCTGATTAGCTAGTGGGTAAGGCGCATTCCTTACCAAGGCGATGATGATGCTCTGA 789

QY 268 GAGGAGCAACGAGCAACTGGGACTGAGCAACGCGCCAGACTCTTACGAGAGGACAGT 327  
 DB 788 AGGAGTGAATCAACCACTGGACTGAGACAGCGCCAGACTCTTACGAGAGGACAGT 729  
 QY 328 GGGGAATTTTGGACAATGGCGCAAGCCTGATCCAGCAATGCGCGTGAAGTGAAGGC 387  
 DB 728 GGGGAATTTTCCGCAATGGCGCAAGCCTGATCCAGCAATGCGCGTGAAGTGAAGGC 669  
 QY 388 CTTGCGGTTGTAAGCTCTTTCAGTGCAGAGAGAAAGTTACGGTAATATCGTACCC 447  
 DB 668 CTACGGGTCCTGAACCTTCTTCCAGAGAGAA-----AGCA 632  
 QY 448 ATGACGGTATGACAGAAAGACCGGCTATCTAGTGCAGACCGCGGTAATACGT 507  
 DB 631 ATGACGGTATCTGGGAAATAGCATGGCTAATCTGTGCAGACCGCGGTAATACAG 572  
 QY 508 AGGGTGCAAGCCTTATCGGAATTAATGCGGCTGAAGAGGTGCGCAGCGCGCTTGAAGT 567  
 DB 571 AGGATGCAAGCCTTATCGGAATTAATGCGGCTGAAGAGGTGCGTGTAGGTGCTTTTAAGT 512  
 QY 568 CAGATGGAATCCCGGGCTTAACCTGGGAATTGCGTTGAATTAACAACTAGAGT 627  
 DB 511 CCGCGCTCAATCCAGGGCTCAACCTGAGCAGCGCGGTGAACCTACCAAGCTTGAAGTA 452  
 QY 628 TGGCAGAGGAGAGTGGAAATTCATGTGTAGCAGTGAATGCGTAGAGATATGAGAAACA 687  
 DB 451 CCGTGAAGGAGAGGAAATTCGAGTGAAGCGGTGAATGCGTAGAGATATGAGAAACA 392  
 QY 688 TCGATGGCGAAGGAGCGCTCTGAGGTGAACATGACGCTCATGCAAGAAAGCGTGGAG 747  
 DB 391 CCAAGCGGAAAGCACTGCTGAGGCGCACTGACATGACATGAGAGAGCAAGGCTAGGGGAG 332  
 QY 748 CAACAGATTTGATACCTGTGTAGTCCAGCGCTTAACGATGTAATCTAGTTTGGGC 807  
 DB 331 CCAATGGGATTTGATACCTGTGTAGTCCAGCGCTTAACGATGTAATCTAGTTTGGGC 272  
 QY 808 CTTATTAAGCTTG--GTAAAGCAAGCTTAACGCGTGAAGTTGACCGCTGGGAGTACGTC 865  
 DB 271 GTATGACCCCGTGCAGTGTGTAGTGAAGCTTAAGTATCCGCTGGGAGTACGTC 212  
 QY 866 GCAAGATTTAACTCAAGAAATTAAGCGGAGCCGCAAGCGGTGATTAATGATGAT 925  
 DB 211 GCAAGATTTAACTCAAGAAATTAAGCGGAGCCGCAAGCGGTGATTAATGATGAT 152  
 QY 926 AATTGATGCAACGGAAGAACTTACCTTACCTTGAACATGTAGCGAATTTTCTAGAGAT 985  
 DB 151 AATTGATGCAAGGAGAAAGAACTTACCTTGAACATGTAGCGAATTTTCTAGAGAT 92  
 QY 986 AATTATAGT--CTTGGGAAAGCTTAACAGAGTGTGCAATGCTGTGCTAGCTCGATGTC 1044  
 DB 91 AAGAGGGTGCCTTTCGGGAAAGCGGACACAGGTGTGATGCTGTGCTAGCTCGATGTC 32  
 QY 1045 TGAGATGTTGGGTTAAGTCCCGCAACGAGCG 1075  
 DB 31 TAAAGTGTGGGTTAAGTCCCGCAACGAGCG 1

RESULT 51  
 BH535640/c 848 bp DNA linear GSS 14-DEC-2001  
 LOCUS BOGYL94TF BOGY Brassica oleracea genomic clone BOGYL94, genomic  
 DEFINITION survey sequence.  
 ACCESSION BH535640  
 VERSION BH535640.1 GI:17766867  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eustoids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 848)  
 Ayele,M., Haeß,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,

TITLE Uterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
 Whole genome shotgun sequencing of *Brassica oleracea* and its  
 application to gene discovery and annotation in Arabidopsis  
 JOURNAL Genome Res. 15 (4), 487-495 (2005)  
 PUBMED 15805490  
 COMMENT Other\_GSSs: BOGYL94TR  
 Contact: Chris Town

TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TP  
 Class: sheared ends.

FEATURES  
 source location/Qualifiers

1..848  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO100DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOGYL94"  
 /clone\_lib="BOGY"  
 /note="Vector: pHD51; Site 1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHD51 using BstXI linkers"

## ORIGIN

Query Match 35.5%; Score 516.6; DB 9; Length 848;  
 Best Local Similarity 77.9%; Pred. No. 8.3e-148;  
 Matches 661; Conservative 0; Mismatches 164; Indels 4; Gaps 3;

449 TGAAGGATGACAGAAAGACACCGGCTAACTACGTCCAGACCGCGGTAAATACGTA 508  
 848 TGACGGATCTCGGGAATAGCATCGCTAACTCTGCGCAGACCGCGGTAAATACGA 789  
 509 GGGTGCAGCGTTAATGGGAATTACTGGGCTTAAGGGTGGCGAGCGCGCTGTAGTC 568  
 788 GGATGCAGAGCGTTATCCGGAATGATTTGGCGTAAAGCGTCTGTA-GTGGGCTTTAAAGTC 730  
 569 AGATGTGAATCCCGGGCTTAACCTGGGAATTCGTTTGAATACAAAGCTAGAGTGT 628  
 729 CGCGTCAATCCAGGGCTTACCTTGGACAGCGCGGTGAACCTTACCAAGCTTGAATAC 670  
 629 GCGAGAGGAGGTGAATTCATGTGTAGCAGTGAATTCGCTAGAGATATGAAAGAAAT 688  
 669 GGTAGGGGAGAGGGAATTTCCGGTGGAGCGGTGAATGCTAGAGATCGGAAGAAACAC 610  
 689 CGATGGGAGAGCGACCTCTCGGGTAACTGACGCTCATGACAGAAAGGTGGGAGC 748  
 609 CAACGGGGAAGCACTGCTGGGCGGACACTGACACTGAGAGAGAAAGCTAGGGAGC 550  
 749 AAACAGATTAGATACCTGTAGTCAACGCGCCCTAAACGATGTAACATGATGTTGGGCC 808  
 549 GAATGGATTAGATACCCAGTAGTCTAGCGGTAAACGATGATCTAGGCGCTGTGCG 490  
 809 TTATTAGGCTTG--GTAAAGAGCTAACGCGTGAAGTTGACCGGCTGGGAGTAGCGGTG 866  
 489 TATGACCCCGTGCAGTGTGTAGCTAACGCGTTAAGTATCCGCTGGGAGTAGTGTGCG 430  
 867 CAAGATTAAACTCAAGGAATTGACGGGAGCCGACACAGCGGTGATTAATGTGAGATTA 926  
 429 CAAGAATGAACCTCAAGGAATTGACGGGAGCCGACACAGCGGTGATTAATGTGAGATTA 370  
 927 ATTGATGCAACGGGAAAAACCTTACCTACCTTGAACATGATGAGGAATTTTCTAGAGATA 986  
 369 ATTGATGCAACGGGAAAAACCTTACCTACCTTGAACATGATGAGGAATTTTCTAGAGATA 310  
 987 GATTAGTTC--CTTCGGGAGCGCTAACACAGTGTGTCATGCTGTCTGTCAGTCTGTGTCGT 1045  
 309 GAGGGGTGCTTCGGGAGCGGAGACAGAGTGTGTCATGCTGTCTGTCAGTCTGTGTCGT 250  
 1046 GAGATGTTGGTTAAGTCCCGAAGAGCGCAACCTTTGATTAATGTCATCTTTGG 1105

Db 249 AAGGTGTGGGTTAAGTCCCGAAGAGCGCAACCTCGTGTTAATGCCACCGTTGAG 190  
 Qy 1106 TTGGGACCTTTATGACTGCCCGTGAACAACCGGAGAGGTGGGATGACGTCAAGT 1165  
 Db 189 TTTGGAAACCTGAACAGACTGCGGTGATAGCCGAGAGAGGTGAGATGACGTCAAGT 130  
 Qy 1166 CCTCATGGCCCTTATGGTATGAGCTTACACACGTATATACATGAGCGCGTTCAGAGGGTTCG 1225  
 Db 129 CATCATGCCCTTTATGCTCCCTGGGAGACACAGCTGTCAATAGCCGGACAAAGGCTCGC 70  
 Qy 1226 CAACCCGAGAGGGGAGCTAATCTCAAAAGCGCGTGTAGTCCGATGGAGTCTGCAC 1285  
 Db 69 GATCCCGAGAGGTGAGCTAATCTCAAAACCGCTCTGAGTTCGATGACAGCTGCAC 10  
 Qy 1286 CTCGACTCC 1294  
 Db 9 CTCGCTGC 1

RESULT 52  
 BH578853/c 857 bp DNA linear GSS 14-DEC-2001  
 LOCUS BOHQ24TR BOHQ Brassica oleracea genomic clone BOHQ24, genomic  
 DEFINITION survey sequence.  
 ACCESSION BH578853.1 GI:117831134  
 VERSION BH578853  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 857)  
 Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
 Uterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
 Whole genome shotgun sequencing of *Brassica oleracea* and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 PUBMED 15805490  
 COMMENT Other\_GSSs: BOHQ24TF  
 Contact: Chris Town

TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.

FEATURES  
 source location/Qualifiers

1..857  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO100DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOHQ24"  
 /clone\_lib="BOHQ"  
 /note="Vector: pHD51; Site 1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHD51 using BstXI linkers"

## ORIGIN

Query Match 35.5%; Score 516.6; DB 9; Length 857;  
 Best Local Similarity 77.9%; Pred. No. 8.3e-148;  
 Matches 671; Conservative 0; Mismatches 164; Indels 26; Gaps 3;

228 GTTAAGGCTTACCAAGGAGCGATGATGTTGCTGAGAGAGACGACGACACTG 287  
 Db 857 GGCATATGCTTACCAAGGAGCGATGATGATGCTGTTCCGAGAGATATGACCACTG 798  
 Qy 288 GGACTGAGACAGCGCCAGACTCTTACCGGAGGACAGCAGTGGGAATTTTGAATATGG 347  
 Db 797 GGACTGAGACAGCGCCAGACTCTTACCGGAGGACAGCAGTGGGAATTTTCCCAATGG 738



Oy		348	GGCAAGCCTTATCCAGCAATGCGGGTAGTGAAAGAAGGCCCTTCGGGTATTGAAGACTTT	407
Db		737	CGAAGCCTTGACGAGCAATGCCGGTGTGAGATGAAAGSCTTACCGGCTCTGAACTTCTT	678
Oy		408	TCAGTCGAGAAAGAAAAGTTACGGTAATAATATCTGTACCATGACGGTATCGACAAGAA	467
Db		677	TTCCACAGAGA-----ACCAATGACGGTATCTGGGGATATA	641
Oy		468	AGCACCGGCTTAATACGTGCAGACGCGCGGTATACTAGTGGGCGCAAGGGTTAATCCG	527
Db		640	AGCATTCGGCTAATCTGTGCCAGACGCGCGGTAAATACAGAGGATGCAAGGTTATCCGG	581
Oy		528	AATTACTGGGCGTAAAGGGTGGCGACGGCGCTTGTAACTAGATGATGTAATATCCCGGGC	587
Db		580	AATGATTGGCGTTAAGCGCTCTGTAGTGGCTTTTAACTCCCGCTCAAAATCCAGGGC	521
Oy		588	TTAACCTTGGGAATTCGTTTGAAACTAACAAAGCTAGAGTGTGGCGAGAGGAGTGAATT	647
Db		520	TCAACCTTGGACAGCGCGGTGGAAATTAACAAAGCTTAGTAAGTGTGGGGCAGAGGAATT	461
Oy		648	CCAATGTAGACAGTGAATTCGTAGATATATGAAAGAACATCGATGGCGAAAGCAGCTTC	707
Db		460	TCCGCTGAGCGGTGAATAATGCGTGAAGATCGGAAAGAACACCAACGGGGAAGCACTCTG	401
Oy		708	CTGGGTTAAACCTGACGCTCAATGCACGAAAGCGTGGGAGCAACAGAGATTAGATACCTT	767
Db		400	CTGGGCGCACTGACACTGAGACAGAAAGCTAGGGAGACGGAATGGAGTTAGATCCCC	341
Oy		768	GGTAGTCCACGCCCTTAAACGATGTCTMACTAGTTGTTGGGCTTATTAGGCTTG--GTAA	825
Db		340	AGTAGTCTTAAGCCGTAAACGATGATCTAAGGCGGTGTGCGGTATGACCCCGGAGTGTCT	281
Oy		826	GAAGCTAACCGCGTGAAGTTGACCGCTTGGGAGATACGGTCCGACGATTAAAACTCAAAG	885
Db		280	GTAGCTAACCGGTTAAAGTATCCCGCTGGGAGTACGTTCCAGAAAGAACTCAAAAG	221
Oy		886	AATTGACGGGGAACCCGCAACGCGGTGAATTATGTGAATTAATTCGATGCAACGGCAAA	945
Db		220	AATTGACCGGGGACCCGCAACGCGGTGAAGCATGTGGTTAAATTCGATGCAAGGAAAGA	161
Oy		946	ACCTTACCTTACCTTGATCATGTAGAGGAATTTTCTAGATATGATTAAGTG-CTTCGGGAA	1004
Db		160	ACCTTACCAAGGCTTGACATGCCGGATTCCTCTTGAAAGAGAGGGGTGCTTCGGGAA	101
Oy		1005	GCTAACACAGAGTGTGCATGTGCTGTCACTGTGTCTGTGTGAGATGTTGGTTTAAGTCC	1064
Db		100	GCGGACACAGAGTGTGCATGTGCTGTCACTGTGTCTGTGTGAGATGTTGGTTTAAGTCC	41
Oy		1065	CGCAACGAGCGCAACCTTGT	1085
Db		40	CGCAACGAGCGCAACCTTGT	20
RESULT_53				
BH473870/c				
LOCUS	BH473870	845 bp	DNA	linear GSS 13-DEC-2001
DEFINITION	BOHQQ2LTR BOHG Brassica oleracea genomic clone BOHQQ27, genomic survey sequence.			
ACCESSION	BH473870			
VERSION	BH473870.1	GI:17681981		
KEYWORDS	GSS:			
SOURCE	Brassica oleracea			
ORGANISM	Brassica oleracea			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; euroside II; Brassicales; Brassicaceae; Brassica.			
AUTHORS	1 (bases 1 to 845) Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Uteirback, T.R., Wortman, J.R., White, O.R. and Town, C.D.			
TITLE	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res.			
JOURNAL	Vol. 15 (4), 487-495 (2005)			
PUBMED	15805490			

COMMENT									
Other GSSES: BOHQZ7TF									
Contact: Chris Town									
TIGR									
9712 Medical Center Drive, Rockville, MD 20850, USA.									
Tel: 301-838-3523									
Fax: 301-838-0208									
Email: cdtown@tigr.org									
DNA is from a doubled haploid provided by Tom Osborn.									
Seq primer: TR									
Class: sheared ends.									
FEATURES									
Source Location/Qualifiers									
1..845									
/organism="Brassica oleracea"									
/mol type="genomic DNA"									
/strain="TO100DH3"									
/db_xref="taxon:3712"									
/clone="BOHQZ7"									
/clone_id="BOHG"									
/note="Vector: pHSOI, site_1: BclXI; 2-3 kb sheared genomic DNA inserted into pHSOI using BclXI linkers"									
ORIGIN									
Query Match 35.4%; Score 516.4; DB 9; Length 845;									
Best Local Similarity 78.1%; Pred. No. 9.6e-148;									
Matches 659; Conservative 0; Mismatches 181; Indels 4; Gaps 3									
Oy	588	TTAACCTGGGAAATTGCCGTTGAACCTCAAGAGCTAGAGTGCGCAGAGGAGGTGAAT	647						
Db	845	TCAACTGGGACAGCGCGGTGAAAATCTACAAGCTGAATCTAGTAGGGGCAGAGGAATT	786						
Oy	648	CCATGTGAGCAGTGAATAATCGTGAAGATATGSAAGAATCATGATGCGAAGGCAGCCTC	707						
Db	785	CCGGGTGAGCGCGTGAATAATGCGTGAAGATCGGAAGAATCACCAACCGGAAAGCATCTCG	726						
Oy	708	CTGGGTTAACTGACGCTCATGACGAGAACCGTGGGGAGCAAACAGATTGATACCT	767						
Db	725	CTGGGCGGACACTGACACTGAGAGACGAACCTAGGGGAGCGAATGGGATTAGATACC	666						
Oy	768	GGTAGTCACGCCCTTAACCATGTATCACTACTGTGTGGGCGCTTATTAGGCTTG--GTAC	825						
Db	665	AGTAGTCCTTACCGGTAAACGATGATCTAGAGCGCTGTGCGTATCGACCCGTGCAGTGCT	606						
Oy	826	GAAGCTAACGGGTGAGTGTGACCGCGCTGGGAGTACGSGTCGCAAGATTAAACTCAAG	885						
Db	605	GTAAGTAAACGGTTAAGTATCCCGCTGGGAGTACGTTCCGMAAATGAACCTCAAG	546						
Oy	886	AATTGACGGGAGCCCGCAACAGCGGTGATTATGTGATTATTCATGCAACGCGAANA	945						
Db	545	AATTGACGGGGGCCCGCAACAGCGGTGAGCATGTGTTTAATTGATGCAAGACGAAGA	486						
Oy	946	ACCTTACCTACCTTTGACATGTAGCGAATTTTCTAGAAATGATGATTAGTG-CITTCGGGAAC	100-						
Db	485	ACCTTACCAAGGGCTTGAATCATCCCGGAATCCTCTTGAAGAAGAGGGGTGCTTCGGGAAC	426						
Oy	1005	GCTAACACAGGTGCGATGCGTGTGCTGACGTGCTGCGAGATGTTGGGTTAAGTCC	106						
Db	425	GCGAACAAGGTGTGATGCGTGTGCTGCTGCAAGTGTGCTGCGGTTAAGTCC	366						
Oy	1065	CGAACAGAGCCCAACCCCTTGTATTAATTGCGATCATTTGGTTGGGCACTTTAATGAGC	112						
Db	365	CGAACAGAGCCCAACCCCTGTGTTAAGTTGTCACCGTTGAAGTTGGAACCTCGAACAGNC	306						
Oy	1125	TGCGGTGACAAACCGAGAGAAAGTTGGGATGACGTCAAGTCTCTATGCGCTTATGGGT	118						
Db	305	TGCGGTGATTAAGCCCGAGAAAGGTGAAGATGACGTCAAGTCATCATGCCCCCTTATGCC	246						
Oy	1185	AGGCTTCAACGTAAATCAATGGCGCGTACAGAGGGTTGCCAACCCCGAGGGGAGGT	124						
Db	245	TGGGCGACACGTCATCAATGGCGGGAACAAAGGTCGCGATCCCGAGGGGTAGCT	186						
Oy	1245	AATCTCAGAAAGCGGTGTAGTCCGGAATCGAAGTCTCAACTCGACTCCGTGAAGTCGG	130						
Db	185	AATCTCAAAAACCGTCTCTAGTTGGAATTCGAGGCTCAACTCGCTCTGACGAAGCCGG	126						



Qy	1305	AATGCGTAGTATCGCGATGATAG-CATGTCGGGGGAATATGATTCGCCGGTCTTGTAAC	1363
Db	125	AATGCTATGTAATGCCCGGTGACCACTACGGCGGTAATTCGTCGCCGGCTTGTAAAC	66
Qy	1364	ACCGCCCTCAACACCATGGAGTGGTTTACACAGAACAGGTACTTAACCGTAAGAG	1423
Db	65	ACCGCCCTCACTATGAGAGCTGGCATGCCAGAACTGTTACCTTAACCGCAAGAG	6
Qy	1424	GGCG 1427	
Db	5	GGGG 2	
RESULT 54			
B2501502			
LOCUS	B2501502	831 bp	DNA linear
DEFINITION	BONDES7TR BO.1.6.2_KB_tot	Brassica oleracea	genomic clone BONDE57,
ACCESSION	B2501502		genomic survey sequence.
VERSION	B2501502		
KEYWORDS	B2501502.1	GI:27018323	
SOURCE	GS5.		
ORGANISM	Brassica oleracea		
REFERENCE	Brassica oleracea		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
TITLE	Whole genome shotgun sequencing of Brassica oleracea and its		
JOURNAL	application to gene discovery and annotation in Arabidopsis		
PUBMED	Genome Res. 15 (4), 487-495 (2005)		
COMMENT	15805490		
	Other_GSSs: BONDES7TR		
	Contact: Chris Town		
	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA.		
	Tel: 301-838-3523		
	Fax: 301-838-0208		
	Email: cdtown@tigr.org		
	DNA is from a doubled haploid provided by Tom Osborn.		
	Seq primer: TR		
	Class: sheared ends.		
FEATURES	Location/Qualifiers		
source	1..831		
	/organism="Brassica oleracea"		
	/mol_type="genomic DNA"		
	/strain="TO100DH3"		
	/db_xref="taxon:37112"		
	/clone="BONDE57"		
	/clone_1fb="BO.1.6.2_KB_tot"		
	/note="Vector: pHD1; Site 1: BclXI; 1.6-2 kb sheared		
	total DNA inserted into pHD1 using BclXI linkers"		
ORIGIN			
Query Match	35.4%;	Score 516.2;	DB 9;
Best Local Similarity	78.1%;	Pred. No. 1.1e-147;	
Matches	646;	Conservative 0;	Mismatches 178;
		Indels 3;	Gaps 2
Qy	463	GAAGAGCAGCGGCTTAAGTACGTCGCCAGACCGCGCGTAATACGTAGGGTGCAGCGTTA	522
Db	5	GGAATAGCATCGGCTTAAGTACGTCGCCAGACCGCGCGTAATACGTAGGGTGCAGCGTTA	64
Qy	523	ATCGGAATTAATGCGGCTTAAGTACGTCGCCAGACCGCGCGTTAGTACGATGTGAATCCG	582
Db	65	TCCGGAATTAATGCGGCTTAAGTACGTCGCCAGACCGCGCGTTAGTACGATGTGAATCCG	124
Qy	583	CGGCTTAATGCGGCTTAAGTACGTCGCCAGACCGCGCGTTAGTACGATGTGAATCCG	642
Db	125	AGGCTTAATGCGGCTTAAGTACGTCGCCAGACCGCGCGTTAGTACGATGTGAATCCG	184
Qy	643	GAATTCATGTGTAGCATGTAATGCGGTGAGATATGAGAAGAACATCATGCGCAAGGCA	702

Db	185	GAAATTTCGGTGGAGCCGTGAATTGCGTAGATCGGAAAGAACCAACGGCGAAAGCA	244		
OY	703	GCCTCTGGGTGTTAACACTGACGCTCATGACAGAAAGCGTGGAGCAACACGATTAGAT	762		
Db	245	CTCTGCTGGGCGCACATGACACTGAGAGACGAAAGCTGAGGGAGCAATGGGATTAGAT	304		
OY	763	ACCTGTGATGCCACGCCCTTAAACGATGTCACTAGTGTGGGCTTATTAGGCTTG--	820		
Db	305	ACCCCACTAGTCCCTAGCCGTAAACGAATGATCTAGGCGCTGTGCGATCCGCGCA	364		
OY	821	GTAACGAAGCTAACGCGTGAAGTTGACCGGCTGGGAGTAGACGTCGCAAGATTAAATC	880		
Db	365	GTGCTGATGCTAACGCGCTTAAATGATATCCCGCTGGGAGTAGCTTCGCAAGATGAATC	424		
OY	881	AAAGAAATTGACGCGGAGCCGCGACACAGCGGTGATTTATGTGATTTAATTCGATCAACGC	940		
Db	425	AAAGAAATTGACGCGGAGCCGCGACACAGCGGTGATTTATGTGATTTAATTCGATCAACGC	484		
OY	941	GAAAAACCTTACTCTACCTTGCATGTAGCGAATTTTCTAGAGATTAATTAATGTC--CTTCG	999		
Db	485	GAAAGAACTTAAACGAGGCTTGAACATGCGCGGAATCTCTTGAAGAGAGGGGTGCTTCG	544		
OY	1000	GGAACGCTTACACAGGTGCTGACAGTGTGCTGACGTCGTCGTGATGATGTTGGGTTA	1059		
Db	545	GGAACGCGGACACAGGTGCTGACAGTGTGCTGACGTCGTCGTGATGATGTTGGGTTA	604		
OY	1060	AGTCCCGCAACGAGCGCAACCTTGTGATTAATTTGCCATCTTTGTGGGCACTTTAAT	1119		
Db	605	AGTCCCGCAACGAGCGCAACCTTGTGATTAATTTGCCATCTTTGTGGGCACTTTAAT	664		
OY	1120	GAGACTGCGGTGACAAACCGGAGAAAGTGGGATGACGTCAAGTCTCTACGTGCGCTTA	1179		
Db	665	CAGACTGCGGTGATTAAGCCGAGAAAGTGAAGATGACGTCAAGTCTCTACGTGCGCTTA	724		
OY	1180	TGGGTAGGGCTTTCACAGTAAATAAATGGGCGGTACAGAGGTTGGCCAAACCGGAGGGG	1239		
Db	725	TGCCCTGGGCGACACAGTGTCTACAAATGGCCGAGCAAAAGGTCGCGATCCCGGAGGGT	784		
OY	1240	GAGCTTAATCTCAGAAAGCGCGTCTGATCCCGATCGAGTCTGCAAC	1286		
Db	785	GAGCTTAATCTCAGAAAGCGCGTCTGATCCCGATCGAGTCTGCAAC	831		
RESULT 55	BH552149	835 bp	DNA	linear	GSS 14-DEC-2001
LOCUS	BH552149				
DEFINITION	BOHFP72TR BOHF Brassica oleracea genomic clone BOHFP72, genomic				
ACCESSION	BH552149				
VERSION	BH552149.1				
KEYWORDS	GSS.				
SOURCE	Brassica oleracea				
ORGANISM	Brassica oleracea				
REFERENCE	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;				
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;				
TITLE	rosidae; euroside II; Brassicales; Brassicaceae; Brassica.				
JOURNAL	Ayle, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,				
PUBMED	Uteerback, T.R., Wortman, J.R., White, O.R. and Town, C.D.				
COMMENT	Whole genome shotgun sequencing of Brassica oleracea and its				
	Application to gene discovery and annotation in Arabidopsis				
	Genome Res. 15 (4), 487-495 (2005)				
	15805490				
	Other GSSs: BOHFP72TF				
	Contact: Chris Town				
	TR				
	9712 Medical Center Drive, Rockville, MD 20850, USA.				
	Tel: 301-838-3523				
	Fax: 301-838-0208				
	Email: cdtown@tigr.org				
	DNA is from a doubled haploid provided by Tom Osborn.				
	Seq primer: TR				

Class: sheared ends.  
Location/Qualifiers  
1..835  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="Tol000H3"  
/db\_xref="taxon:3712"  
/clone\_lib="BOHP2"  
/clone\_id="BOHP"  
/note="Vector: pHS1; Site 1: BstXI; 2-3 kb sheared  
genomic DNA inserted into pHS1 using BstXI linkers"

ORIGIN

Query Match 35.4%; Score 516.2; DB 9; Length 835;  
Best Local Similarity 77.7%; Pred. No. 1.1e-147;  
Matches 649; Conservative 0; Mismatches 183; Indels 3; Gaps 2;

479 ACTACGTCGACGAGCGCGGTAAATACGTAGGGTGAAGGCTTAATCGGAATTACTGGG 538  
1 ACTCTGTCCAGACGCCGCGTAAATACAGAGATGCAAGCTTAATCCGGAATATTTGGG 60  
539 GTAAAGGTCGCGAGCGGCGCTTTGTAAGTGAATGTGAAATCCCGGGCTTAACTGGGA 598  
61 GTAAAGGTCGTAAAGGCTTTTAAATCCCGCTCAATCCAGGGCTCAACCTCGGA 120  
599 ATTGCTTTGAAATCTAAAGCTAGAGTGTGCGAGAGGAGTGGAAATTCATGTGTAGC 658  
121 CAGGCGGTGAAACTCAAGCTTGTAGTACGTAGGGGCAAGGGAATTTCCGGTGAAGC 180  
659 AGTGAATGGGTAGATATGAAGAACTGATGGCGAAGGAGCTCTCGGTTTAA 718  
181 GGTGAATGGGTAGATATGAAGAACTGATGGCGAAGGAGCTCTCGTGGCCGACA 240  
719 CTGACGCTCATGACGAAAGCGTGGGAGGAAACAGGATTGATACCTGTGATCCAG 778  
241 CTGACACTGAGAGCAAGAACTAGGGGAGCAATGGGATTAGATACCCAGTAGTCTAG 300  
779 CCTTAACGATGTCACTAATTTGGGCTTTATAGGCTTG--GTAAACAACCTAAACG 836  
301 CCGTAAACGATGTCACTAATTTGGGCTTTATAGGCTTG--GTAAACAACCTAAACG 360  
837 GTGAAGTTGACCGCGTGGGAGTACCGGTGCGAAGATTAAACTCAAGGAATTGACGGG 896  
361 GTTAAGTATCCCGCTGGGAGTACGTTCCCAAGATGAATCAAGGAATTGACGGG 420  
897 ACCCGCAAGCGGTGATTAATGTGAATTAATTCATGCAACGCGAATACTTACCTAC 956  
421 GCCCGCAAGCGGTGATTAATGTGAATTAATTCATGCAACGCGAATACTTACCTAC 480  
957 CTTGACATTAAGCAATTTTCTAGAGATGATTAGTG--CTTGGGAAACCTTAACACAG 1015  
481 GCTTGAATGCGCGGATCTCTTGAAGAGAGGGGCTTTCGGGAACCGGACACAG 540  
1016 TGTGATGATGCTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075  
541 TGTGATGATGCTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
1076 CAACCTTGTCAATTAATTTGCAATTAATTTGCTGGGCACTTAAATGAGATGCGGTGAC 1135  
601 CAACCTTGTCAATTAATTTGCAATTAATTTGCTGGGCACTTAAATGAGATGCGGTGAC 660  
1136 AACCGGAGAAAGTGGGAGTGAAGTCAAGTCTCATATGCTTATGGTATGGGCTTACA 1195  
661 AGCGGAGAAAGTGGGAGTGAAGTCAAGTCTCATATGCTTATGGTATGGGCTTACA 720  
1196 CGTAAATACATGGCGGTACAGAGGTTGCAACCGCGAGGGGAGAGCTATCTCAGAAA 1255  
721 CGTGTCAATATGGCGGTACAGAGGTTGCAACCGCGAGGGGAGAGCTATCTCAGAAA 780  
1256 GCGCGTGTATGCGGATCGGATCTGCAACTCGACTCCGTGAAGTGGAGATGCG 1310  
781 CCGGTCTCTAGTTGCGATTCAGAGCTGCAACCTGCTGATGAAGCGGATGCG 835

RESULT 56  
CWM01852 1050 bp DNA linear GSS 23-SEP-2004  
LOCUS ZMBH90002h20.f ZMBH9 Zea mays genomic clone ZMBH90002h20 5',  
DEFINITION genomic survey sequence.  
ACCESSION CWM01852  
VERSION CWM01852.1 GI:52576905  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 1050)  
Ma, J., SanMiguel, P., Liu, R., Haller, K., Soderlund, C. and  
Bennetzen, J.  
ZMBH sequences  
Unpublished (2004)  
TITLE ZMBH sequences  
JOURNAL Unpublished (2004)  
COMMENT Contact: Jeff Bennetzen  
Bennetzen Lab  
The University of Georgia  
Department of Genetics, C426a Life Sciences Building, Athens, GA  
30602, USA  
Tel: 706-542-3698  
Fax: 706-583-0972  
Email: matre@uga.edu  
Plate: 0002 Row: h Column: 20  
Classes: BAC ends.

FEATURES  
source Location/Qualifiers  
1..1050  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMBH90002h20"  
/tissue\_type="immature ear"  
/dev\_stage="6-8 weeks"  
/lab\_host="DH10B"  
/clone\_lib="ZMBH9"  
/note="Vector: TOPopcr4; Site\_1: EcoRI; Site\_2: EcoRI"

ORIGIN

Query Match 35.4%; Score 515.6; DB 10; Length 1050;  
Best Local Similarity 75.4%; Pred. No. 1.8e-147;  
Matches 706; Conservative 0; Mismatches 224; Indels 6; Gaps 5;

527 GAATTACTGGGCGTAAAGGCTGCGAGCGGCTTGTAGTGAATGTAATCCCGGG 586  
2 GGATGATTTGGGCTAAAGCGCTGTAGGTGGCTTTTCAAGTCCGCGTCAATCCAGGG 61  
587 GTTAACCTGGGAATTCGCTTTGAAACTCAAAAGCTAGAGTGTGCGAGAGGAGTGAAT 646  
62 CTCAACCTCTGACAGCGCGGTGAAAATCAAGCTGAGTACGTAGGGGCGAGAGGAAT 121  
647 TCCATGTGACAGGAATTCGCTAGATATGAAGAACTATGCGAGAGGAGCGCT 706  
122 TTCGGGTGAGCGGTAAATTCATTGATCGAAAGAACACACGCGGAGAACCTT 181  
707 CTTGGGTTAACTGACGCTCATGACGAAACGTTGGGAGGCAACAGGATTAGTATCC 766  
182 GCTGGGCGGACACTGACATGAGAGCGAAAGCTAGGGGAGCAAAATGGGATTAGAGACC 241  
767 TGTAGTCAAGCGCTTAACAGATGTCACTAGTGTGGGCTTATTAAGCTTG--GTAA 824  
242 CAGTAGTCTTAACCGTAAACGATGATTAAGTGTGCTGTGCACTGACCCGTGACATGC 301  
825 CGAAGCTAACGCTGAAGTTGACCGCTGGGAGTACCGTGCAGAGTTAAACTCAAG 884  
302 TGTAGCTAACGCTGAAGTTGACCGCTGGGAGTACCGTGCAGAGTTAAACTCAAG 361  
885 GAATTAACGCGGAGCCGCAACAGCGGTGATTAATGATTAATTCATGACGAGAA 944

D	b		362	GAAATTGACGGGGGCCCCCAACAAGCGGTGGAGCATTGTGGTTTAATTCATGAAGAAGCAAG	421
Oy			945	AACCTTACCTAACCCCTTGACATGTAGCCGAATTTCTTAGAGATAGATTAGTG-CTTCGGGAA	1003
D	b		422	AACCTTACGAGGGCTTGACATGCGCGCAATCCTCTTAGAAGAAGAGGGGTGCCTCCGGAA	481
Oy			1004	CGCTTAACACAGTGTCTCATATGGCTGTGTGACCTCGTGTGCTGTGATGATTGGGTTAAGTC	1063
D	b		482	CGCGACAACAGTGTGTGATGAGTGGCTGTGCTGACCTGTGCCTTAAGTGTGTGGTTAAGTC	541
Oy			1064	CCGCAACGAGCGCAACCCCTTGCTCATTTAATATGSCATCATTTGGTTGGGCACTTTAATAGA	1123
D	b		542	TCCGACGAGCGCAACCTCGTGTGTTAGTTGCCA-CATAGATTGGAAACCCTGAACGA	600
Oy			1124	CTGCCGCTGACMAACCGAGGAGAGGTGGAGATGACGTCAAGTCTCATGSCCTTTATGG	1183
D	b		601	CCGCGGCTGTTAAGCCCGAGGAGAGAGATGAGGCAATCATCATGCCCTTATAGCC	660
Oy			1184	TAGGGCTTCAACATGTAATACAAATGGCGGTTACAGAGGTTTCCAAACCCGAGAGGGGAGC	1243
D	b		661	CTGGGCGCACACACGCTTACAAATGGGCGGACAAAGGCTCGCATCTGCGAGAGGTGAGC	720
Oy			1244	TAACTCGAAGAAAGCGCTGTAGTCCGAGTCGAGTGTGCAACTGCACTCCGTAAGTGC	1303
D	b		721	TAACTCCAAAACCGTCTCTCAAGTTGCAAGTTGCAAGCTGCAACTGCTGATGAAGAG	780
Oy			1304	GAATGCTAGTAAATGCGCGATCAG-CATGTGCGGCTGAATACTGTTCCGGGCTTTGACA	1362
D	b		781	GAATGCTAGTAAATGCGCGATCAGCATACGCGGCTGAATCCGTTCCGGGCTTTGTACA	840
Oy			1363	CACGCCCCGTCAACATGAGAGTGGGTTTCAACGAGACAGTAGTC-TAAACGTAGG	1421
D	b		841	CACCCCGCTGCACATATAGAGCTGGCCAGGTTTGAAGTCATTACCTTAAACGTAAAG	900
Oy			1422	AGGGGCTTCCACAGGTGAGATTCACTGACTGGGGTG	1457
D	b		901	AGGGGATGCTTAAGGCTTAGGCTTGCCGACTGGAGTG	936
<hr/>					
RESULT 57					
BH704946/c					
LOCUS BOMLQ71TF_BO_2_3_KB Brassica oleracea genomic clone BOMLQ71,					
DEFINITION genomic survey sequence.					
ACCESSION BH704946					
VERSION BH704946.1 GI:18787033					
KEYWORDS GSS.					
SOURCE Brassica oleracea					
ORGANISM Brassica oleracea					
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;					
TITLE rosoids; eurosids II; Brassicales; Brassicaceae; Brassica.					
1 (bases 1 to 835)					
AYEYE,N., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,					
UTTERBACK,T.R., Wortman,J.R., White,O.R. and Town,C.D.					
Whole genome shotgun sequencing of Brassica oleracea and its					
JOURNAL application to gene discovery and annotation in Arabidopsis					
PUBMED Genome Res. 15 (4), 487-495 (2005)					
15805480					
COMMENT Other_GSSes: BOMLQ71TR					
Contact: Chris Town					
TIGR					
9712 Medical Center Drive, Rockville, MD 20850, USA.					
Tel: 301-838-3523					
Fax: 301-838-0208					
Email: cdtown@tigr.org					
DNA is from a doubled haploid provided by Tom Osborn.					
Seg primer: TF					
Class: sheared ends.					
Location/Qualifiers					
1..835					
FEATURES					
source					
/organism="Brassica oleracea"					
/mol_type="genomic DNA"					

```

/strain="T01000DH3"
/db xref="taxon:3712"
/c1one="BOML071"
/c1one.1b="BO_2_3_KB"
/note="vector: pHOSt; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOSt using BstXI linkers"

```

Query Match	35.38	Score 514.8	DB 9	Length 835
Best Local Similarity	78.24	Pred. No. 36-147		
Matches 644	Conservative 0	Mismatches 177	Indels 3	Gaps 2
QY	508	AGGGTGCAGACGGTTAATCGGAATTAATCTGGGCGCTAAAGGGTGCAGACGGCGCTTGTAAAGT	567	
DB	833	AGAGTGCAGACGGTTAATCGGAATTAATCTGGGCGCTAAAGGGTGCAGACGGCGCTTGTAAAGT	774	
QY	568	CAGATGTGAATATCCCGGGCTTAACTTGGGAATTTGGCTTTGAAAATTAACAAAGCTAAGTC	627	
DB	773	CCGCGCTCAATATCCCAAGGGCTCAACCTTCACAGGGCGGTGGAACATCAACAAAGCTTAAAGTA	714	
QY	628	TGCGAGAGGAGAGGTGGAATTCATATGTATACAGTGAATATGCTAAGATATATGGAAGACA	687	
DB	713	CGGTATAGGGCGCAGAGGAATTTCCGGTGTAGACCGGTGAATATCTAAGATATGGAAGACA	654	
QY	688	TCGATGTGCAGAGGAGCGCTCCTGGGTATTAACATGACCGCTCATGCAACGAAAGCGTGGAG	747	
DB	653	CCAAACGGCGAAGACACTCTGCTGGGCGGACACTGACACTGAGAGACGAAAGCTATAGGGAG	594	
QY	748	CAAAACAGATTAGATATCCCTGTATGTCCACGCCCTTAACGATGTCAACTATGTTTGGGC	807	
DB	593	CGAATGGGATTAAGATATCCCAAGTATGCTATAGCCGTAACATGATATCTATGGCGCTGTGC	534	
QY	808	CTTATTTAGGCTTG--GTAAAGAGCTTAACCGCTGAATTTGACCGCCCTGGGAGATAGCGTC	865	
DB	533	GTATTCGACCCCTGTGACGTGTGTACTTAACCGCTTAATATATCCGCTCTGGGAGATACCTTC	474	
QY	866	GCAAGATTAAACTCAAAAGGATTTGACGGGACCCGACAAAGCGGTGATTAATGTGATT	925	
DB	473	GCAAGATTAAACTCAAAAGGATTTGACGGGACCCGACAAAGCGGTGATTAATGTGATT	414	
QY	926	AATTGCATGACGCGAAGAAACCTTACCTACCTTGAATGTAGCGAATTTTCTAGAT	985	
DB	413	AATTGCATGACGCGAAGAAACCTTACCAAGGCTTGAATGACGCGAATCTCTTGAAG	354	
QY	986	AGATTATGTG-CTTGCGGAAACGCTTAACACAGGTGTGATATGGCTGTCTCAGCTCGTGTG	1044	
DB	353	AGAGGGGTGCTTGTGGGAAACGCGACACAGGTGTGATATGGCTGTCTCAGCTCGTGTG	294	
QY	1045	TGAGATGTTTGGGTTAATGTCGCCGAAACGAGCGCAACCTGTATTAATTTGCATATTTG	1104	
DB	293	TAAAGTGTGTTGGTTAATGTCGCCGAAACGAGCGCAACCTGTATTAATTTGCATATTTG	234	
QY	1105	GTTGGGCACTTTAATAGAGACTGCGGTGACAAACCGAGAGAAAGGTGGGATGACGTCAAG	1164	
DB	233	GTTTGAACCTCTGAACAGACTGCGGTGATTAAGCGGAGAGAAAGGTGAGATGACGTCAAG	174	
QY	1165	TCTCATATGACCTTTATGGGTAGGGCTTCAACGTTAATCAATGCGCGGTACAGAGGCTTG	1224	
DB	173	TCATCATATGACCTTTATGCTGTGGGACACACGTTGTTAATATGCGCGGACAAAGGGTGTG	114	
QY	1225	CCAAACCGCGAGGGGAGACTAATCTCAGAAAGCCGCTGTATGTCGGATGGAAGTGTGCA	1284	
DB	113	CGATCCCGCAGGGGTGACTAATCTCAAAAAACCGTCTCAATTCGAGTTTGCAGAGGTGTGCA	54	
QY	1285	ACTGCACCTCGTGAAGCGGAATCGCTATATCGCGGATCAGC	1328	
DB	53	ACTTCGCTGTGATGAAGCCGGAATGCTATATATCGCGGTACG	10	

RESULT 58				
CV291378/c				
LOCUS	CV291378	837 bp	mRNA	linear
DEFINITION	ac001-1mb3-g08	Ac001	Asparagus officinalis	CDNA clone
				EST 08-JUN-2005

ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
af01-1ms3-g08 5', mRNA sequence.									
Cv291378									
Cv291378.1	GI:52577614								
EST.									
Asparagus officinalis (garden asparagus)									
Asparagus officinalis									
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;									
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;									
Asparagaceae; Asparagus.									
1 (bases 1 to 837)									
dePamphilis,C., Carlson,J., Ma,H., Solts,D., Solts,P.,									
Opdenheimer,D., Fronlich,W., Doyle,V., Tansley,S., Webb,M.,									
Leebens-Mack,J., Landherr,L., Ilt,D. and Wall,K.									
Generation of ESTs from early male inflorescences of Asparagus									
officinalis									
unpublished (2004)									
Contact: Claude dePamphilis or James Leebens-Mack									
Mueller Laboratory									
Penn State University									
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn									
State University, University Park, PA 16802, USA									
Tel: 814 863 6413									
Fax: 814 865 9131									
Email: cw43@psu.edu or jh110@psu.edu									
The sequence provided is trimmed of vector and low quality regions.									
Full sequence and original trace file are available from the Plant									
Genome Network website ( <a href="http://pbn.cornell.edu">http://pbn.cornell.edu</a> )									
Plate: af01-1ms3 row: g column: 08									
Seq primer: M13p.									
Location/Qualifiers									
1..837									
/organism="Asparagus officinalis"									
/mol_type="mRNA"									
/db_xref="PGN:af01-1ms3-g08"									
/db_xref="taxon:4686"									
/clone="af01-1ms3-g08"									
/rname="male inflorescences"									
/lab_host="SOUR"									
/clone_id="af01"									
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;									
Site 2: XhoI; This is a directionally cloned,									
non-normalized library. This library has been generated by									
the Floral Genome Project (FGP). The Floral Genome Project									
is funded by NSF's Plant Genome Research Program									
(DBI-0115684). More information about the project can be									
obtained at <a href="http://fgp.bio.psu.edu">http://fgp.bio.psu.edu</a> "									
ORIGIN									
Query Match	35.3%	Score 514.6;	DB 7;	Length 837;					
Best Local Similarity	77.9%	Pred. No. 3.4e-147;							
Matches 669;	Conservative 0;	Mismatches 164;	Indels 26;	Gaps 3;					
221 TTGGTGGGTTAAAGGCTTACCAAGCGCAGATCACTAGTGTCTGAGAGCAGCACGC	280								
837 TTGGTGAGGCAATAGCTTACCAAGCGCATGATCACTAGTGTCTGAGAGATGATCAGC	778								
281 CACACTGGGACTGAGACACCGGCCAGATCTCTTAGCGGAGCGACAGCATGGGGGAATTTTGA	340								
777 CACACTGGGACTGAGACACCGGCCAGATCTCTTAGCGGAGCGACAGCATGGGGGAATTTTCCG	718								
341 CAATGGGCGCAAGCTGATCCAGCAATCCCGCTGATGAGAGAGAGGCTTCCGGTTGTA	400								
717 CAATGGGCGCAAGCTGATCCAGCAATCCCGCTGATGAGAGAGAGGCTTCCGGTTGTA	658								
401 AGCTCTTTCAGTGAGAAAGGTTACGGTAAATAATCGTGACCCATGACGGTATCGA	460								
657 ACTCTTTTTCAGAGAGA-----AGCATGACGGTATCTCG	621								
461 CAGAGAGAGCACCAGCTTAATCACTGCGCAGAGCGCGGTAAATCGTAGGGTGAAGGCT	520								
620 AGGATAAGCATCGGCTTAATCTGTGCGCAGAGCGCGGTAAAGCAGAGAGATGCAAGGT	561								
521 TAATCGAATTAATCGGCGTAAAGGCTGCGAGCGCGCTTGTAAGTCAGATGTGAATC	580								

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	BZ470443/c		BZ470443									
Qy		862 bp DNA	BONRZ51TR_BO_1.6_2_KB_tot									
Db		1near GSS 13-DEC-2002	BONRZ51TR_BO_1.6_2_KB_tot									
Qy		1057	BZ470443									
Db		81	BZ470443									
Qy		997	BZ470443									
Db		201	BZ470443									
Qy		938	BZ470443									
Db		261	BZ470443									
Qy		820	BZ470443									
Db		321	BZ470443									
Qy		760	BZ470443									
Db		381	BZ470443									
Qy		700	BZ470443									
Db		321	BZ470443									
Qy		820	BZ470443									
Db		261	BZ470443									
Qy		938	BZ470443									
Db		201	BZ470443									
Qy		820	BZ470443									
Db		321	BZ470443									
Qy		760	BZ470443									
Db		381	BZ470443									
Qy		700	BZ470443									
Db		321	BZ470443									
Qy		820	BZ470443									
Db		261	BZ470443									
Qy		938	BZ470443									
Db		201	BZ470443									
Qy		820	BZ470443									
Db		321	BZ470443									
Qy		760	BZ470443									
Db		381	BZ470443									
Qy		700	BZ470443									
Db		321	BZ470443									
Qy		820	BZ470443									
Db		261	BZ470443									
Qy		938	BZ470443									
Db		201	BZ470443									
Qy		820	BZ470443									
Db		321	BZ470443									
Qy		760	BZ470443									
Db		381	BZ470443									
Qy		700	BZ470443									
Db		321	BZ470443									
Qy		820	BZ470443		</							

/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONR251"  
/note="vector: pHO51; Site 1: BstXI, 1.6-2 kb sheared  
total DNA inserted into pHO51 using BstXI linkers"

## ORIGIN

Query Match 35.3%; Score 514.4; DB 9; Length 862;  
Best Local Similarity 77.4%; Pred. No. 4e-147;  
Matches 646; Conservative 0; Mismatches 181; Indels 3; Gaps 2;

444 ACCATGACGGTATCGACAGAGAGACCGGTAATCACTAGTCCAGACCGCGGTAA 503  
830 AGCAATGACGGTATCTGGGAAATAGCATCGGCTAATCTGTGCGACAGCCCGTAAT 771  
504 ACCTAGGGTGCAGACGTTAATCGAATTAATCTGGCGCTAAAGGGTGGCAGCGGCTTGT 563  
770 ACAGAGATGCAAGCGTTATCCGGAATGATTTGGGCGTAAAGCGTCTGTAGTGGCTTTT 711  
564 AATTCAGATGTAATCCCGGGCTTAACCTGGGAATGCGTTGAAACTACAAAGCTAG 623  
710 AAGTCGCGCGTCAAAATCCAGGCTCAACCTGAGACAGCGGTGAAACTACCAAGCTTG 651  
624 AGTGGCAGAGGAGGTGGAATTCATGTGTAGCATGTAATGCGTGAATATGAG 683  
650 AGTACGTTAGGGGCGAGAGGAAATTTCCGCTGAGAGCGGTGAATGCGTAAGATCGAAG 591  
684 AACATGATGCGGAGGACGCTCTCGGTTAACTGACGCTCATACGAAAGCTGG 743  
590 AACACCAACGCGGAAACCATCTGCTGGGCGACATGACATGAGAGAGCAAAAGCTAG 531  
744 GAGACAAACAGATTAATTAATCCCTGTATGTCACCGCTTAAAGATGCAACTAGTTGT 803  
530 GAGCGGATGGGATTAATATCCCAAGTATCTAGCCGTAAACGATGATATAGAGGCT 471  
804 GGGCCCTAATAGGCTTGG--GTAAAGAACTAACCGCTGAAGTTGACCGCTGGGAGTAC 861  
470 GTGCGTATCGACCGCTGCAAGTCTGTAGCTAACCGCTTAAGTATCCCGCTGGGAGTAC 411  
862 GGTGCGAAAGTTAAACTCAAAAGGAATTGACGGGAAACCCGCAAGCGGTGGATTATGTG 921  
410 GTTCGCAAGATGAATCACTCAAAAGGAATTGACGGGAGCCCGCAAGCGGTGGAGCATGTG 351  
922 GATTATATTCATGACAGCCGAAACCTTACCTTACCATGTACATGTAGCGAATTTTCTAG 981  
350 GTTAAATTCATGACAAAGCGAAGAACTTACAGAGGCTTGAATGACCGCGAATCTCTTG 291  
982 AGATAGATTAGTG--CTTCGGGAGCGCTAACACAGGTGCTGATGGCTGTCTGACCTGCT 1040  
290 AAAGAGAGGGGTGCTTCGGGAAACCGGACACAGGTGTGATGGCTGTCTGACCTGCT 231  
1041 GTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGGCGCAACCTGTCTCAATTAATGCCATCA 1100  
230 GCGCTAAGGTGTTGGGTTAAGTCCCGCAACGAGGCGCAACCTGTCTCAATTAATGCCATCA 171  
1101 TTTGCTGGGCACTTTTAATGAGATGCGCGGTGACAAACCGGAGAGAAAGGTGGAGTAGCT 1160  
170 TTGAGTTTGGAAACCTGAAACGACTGCGGTGATTAAGCCGAGGAGAGGTGAGTAGCT 111  
1161 CAAATCCCTCATGAGCGCTTAATGAGGCTTACACAGTATATACATGAGCGGTACAGAGG 1220  
110 CAAATCATATGAGCGCTTAATGAGGCTTACACAGTATATACATGAGCGGTACAGAGG 51  
1221 GTTGGCAACCCGCGAGGGGAGCTAATCTCAAAAGCGCGTGTAGTCCG 1270  
50 GTGCGAATCCCGCGAGGGGTGAGCTAATCAAAAACCGTCTCAGTTGG 1

RESULT 60  
B2463792  
LOCUS

B2463792 1030 bp DNA linear GSS 13-DEC-2002

DEFINITION BONKY05TR\_BO\_1.6\_2\_KB\_tot Brassica oleracea genomic clone BONKY05,  
genomic survey sequence.  
ACCESSION B2463792  
VERSION B2463792.1 GI:26749971  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea

REFERENCE 1 (bases 1 to 1030)  
Ayele, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Ulterback, T. R., Wortman, J. R., White, O. R., and Town, C. D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)

## COMMENT

Other GSSs: BONKY05TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Classes: sheared ends.  
Location/Qualifiers

## FEATURES

source

1..1030  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONKY05"  
/note="vector: pHO51; Site 1: BstXI, 1.6-2 kb sheared  
total DNA inserted into pHO51 using BstXI linkers"

## ORIGIN

Query Match 35.3%; Score 513.6; DB 9; Length 1030;  
Best Local Similarity 77.4%; Pred. No. 7.5e-147;  
Matches 661; Conservative 0; Mismatches 189; Indels 4; Gaps 3;

608 GAAACTTCAAAAGCTTATAGTGTGCAAGAGGAGGTGCAATTCATGTGTAGCACTGAATG 667  
2 GAAACTTCAAAAGCTTATAGTGTGCAAGAGGAGGTGCAATTCATGTGTAGCACTGAATG 61  
668 CTTAGAGATTTGGAAGAACTGATGCGGAGGACGCTCTGGGTTAACTGACGCTC 727  
62 CTTAGAGATTTGGAAGAACTGATGCGGAGGACGCTCTGGGTTAACTGACGCTC 121  
728 ATGACGAAAGCGTGGGAGCAAAACAGATTAGTATCCCTGTATGTCACGCGCTTAAAG 787  
122 AGAGACGAAAGCTTATAGGAGGAGGAGGATTTAGTATCCCAATGATCTTACCGCTTAAAG 181  
788 ATGTCATATGTTGGGCTTATTAAGCTTG--GTAAAGAACTTAAAGCGGTGAATG 845  
182 ATGATATCTAGGCGGTGTGCTATGCAACCGGTGAGTGTAGCTTAAACGCTTAAATAT 241  
846 ACCGCTGGGAGATGACGCTGCAAGATTAACTCAAAAGAAATTGACGGGAAACCCGAC 905  
242 CCGGCTGGGAGATGACGCTGCAAGATTAACTCAAAAGAAATTGACGGGAAACCCGAC 301  
906 AGCGTGGATTATGATTAATTCATGACGCGGAAACCTTACCTTACCTTGAAT 965  
302 AGCGTGGATGATGATGATTAATTCATGACGCGGAAACCTTACCTTACCTTGAAT 361  
966 GTAGCGAATTTTCTAGAGATGATTAAGTG--CTTGGGAAAGCTTAAACACAGGTGCTGATG 1024  
362 GCGCGAATCTCTTGAAGAAAGAGGGGTGCTTGGGAAAGCGGACACAGGTGCTGATG 421  
1025 GCTGTCTGATGCTGCTGCTGATGATGTTGGGTTAAGTCCCGCAACGAGGCAACCTTGG 1084

Db 422 GCTGTCGTCAGCTGTCGCGTAAAGTGTGGTTAGTCCCGCAAGAGCAACCTCG 481  
 Qy 1085 TCATTAAATGGCCATCATTTGGTTGGCACTTTAATAGACTGCGGTGACAAACCGGAG 1144  
 Db 482 TGTATTAGTTGCCACCGTTGAGTTGGAAACCTGAAAGACTGCGGTGATTAAGCCCGAG 541  
 Qy 1145 AAGTGGGATGACGTCAGTCTCTCAATGCGCTTAAGGGTAGGGCTTCAACAGTATPAC 1204  
 Db 542 AAGGTAGAGATGACGTCAGTCTCTCAATGCGCTTAAGGGTAGGGCTTCAACAGTATPAC 601  
 Qy 1205 ATGCGCGATGACAGAGGTTGGCCAAACCGCGAGGGAGAGTAAATCTGAAAGCGCGTCT 1264  
 Db 602 ATGCGCGGAGCAAGAGGTGCGCATTCGCGAGAGGTAGCTAATCCAAAGACCGCTCTC 661  
 Qy 1265 ATGCGCGATGACAGTCTGCAACTGCACTCGGTGAAGTGGAAATGCTTAATGCGCAT 1324  
 Db 662 AGTTCCGATTTGACGGCTGCAACTGCGCTGATGAACCGGAAATGCTTAATGCGCGGT 721  
 Qy 1325 CAG-CATGTCGCGGTGAAATACGTTCCCGGTCTTTGACACACCGCGCTGACACCATGG 1383  
 Db 722 CAGCCATACGCGCGGTGAAATGCTTCCCGGTCTTTGACACACCGCGCTGACATGAG 781  
 Qy 1384 AGTGGGTTTCAACGAGAGCAGTGTCTAACCGTAAAGAGGCGCTTGGCCACGCTGAGAT 1443  
 Db 782 AGCTGCCATGCCCCGAAAGTCTTAACTTAAACCGGAGAGGGGGTGGCGAAGCAGGG 841  
 Qy 1444 TCATGACTGGGGTG 1457  
 Db 842 TAGTGACTGAGTG 855

RESULT 61  
 B2459110 1082 bp DNA linear GSS 13-DEC-2002  
 LOCUS B2459110  
 DEFINITION B2459110 BO.1.6.2 KB. tot Brassica oleracea genomic clone BONBR59,  
 genomic survey sequence.  
 ACCESSION B2459110  
 VERSION B2459110.1 GI:26740285  
 KEYWORDS GSS.

SOURCE  
 ORGANISM  
 Brassica oleracea  
 Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 1082)  
 Ayele,M., Haas,B.U., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
 Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 PUBMED  
 COMMENT  
 Other GSSs: BONBR59TF  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.

FEATURES  
 source  
 Location/Qualifiers  
 1..1082  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000H3"  
 /db\_xref="taxon:3712"  
 /clone="BONBR59"  
 /note="Vector: PHOS1; Site 1: BstXI; 1.6-2 kb sheared  
 total DNA inserted into PHOS1 using BstXI linkers"

ORIGIN

Query Match 35.3%; Score 513.6; DB 9; Length 1082;  
 BstXI Local Similarity 73.7%; Pred. No. 7,6e-147;  
 Matches 743; Conservative 0; Mismatches 214; Indels 51; Gaps 5;  
 Qy 3 TGAACGCTGGCGGCAATGCTTTACATATGCAATGTCGAAACGGCAGCAGATGCTTCATCT 62  
 Db 124 TGAACGCTGGCGGCAATGCTTTACATATGCAATGTCGAAACGGCAGCAGATGCTTCATCT 171  
 Qy 63 GGTGCGATGCGGAGACCGGTGATGATGCAATCGGAAGTATCCAGAGAGGGGGTATAC 122  
 Db 172 -GTTTCAGATGCGGAGCGGTGATGATGCAATCGGAAGTATCCAGAGAGGGGGTATAC 230  
 Qy 123 GCATGAAAGATGCTTAATACCGATATCTTAAGAGAGAAACAGGGGATCCAAAGA 182  
 Db 231 AGCTGAAACGGCTGCTTAATACCC--GTAGCTAGAGAGAAAGAGGAATC----- 282  
 Qy 183 CTTTCGCTTTTGAAGCGGCGCATGTCGATTTAGCTAGTGTGGGTAAAGCCTTACCA 242  
 Db 283 -----CGCCGAGAGGGGCTCGCTGTGATTAGCTAGTGTGGGTAAAGCCTTACCA 337  
 Qy 243 AGCGACATCATGATGTTGTTGATGAGAGACGACGACGACCTGAGACACGGC 302  
 Db 338 AGCGATGATCATGATGTTGTTGATGAGAGATGATGATGATGATGATGATGATGATGATG 397  
 Qy 303 CCAGACTCTTAAG 362  
 Db 398 CAGACTCTTAAG 457  
 Qy 363 GCAATGCGCGGTGATGAG 422  
 Db 458 GCAATGCGCGGTGATGAG 515  
 Qy 423 AGTTACGTTAATATATCTGACCCATGACGCTATGACAGAGAGAGAGAGAGAGAGAGAG 482  
 Db 516 -----AGCAATGCGGTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 554  
 Qy 483 CTTGCGAGAGCGCGGTATATAGTATGAGTGCAGACGTTAATCGGAATTAATCTGAGCGTAA 542  
 Db 555 TGTGCGAGAGCGCGGTATATAGTATGAGTGCAGACGTTAATCGGAATTAATCTGAGCGTAA 614  
 Qy 543 AGGTGCGAGCGCGGTATATAGTATGAGTGCAGACGTTAATCGGAATTAATCTGAGCGTAA 602  
 Db 615 AGGTGCGAGCGCGGTATATAGTATGAGTGCAGACGTTAATCGGAATTAATCTGAGCGTAA 674  
 Qy 603 CTTTGAATCTTAAG 662  
 Db 675 CTTTGAATCTTAAG 734  
 Qy 663 AATGCGTGAATGAG 722  
 Db 735 AATGCGTGAATGAG 794  
 Qy 723 CGCTATGACAGAAACGTTGAG 782  
 Db 795 CACTGAG 854  
 Qy 783 AATGATGATCACTAGTATGTTGGCTTATTAAGCTT--GTAGAGAGAGAGAGAGAGAGAG 840  
 Db 855 AATGATGATCACTAGTATGTTGGCTTATTAAGCTT--GTAGAGAGAGAGAGAGAGAGAG 914  
 Qy 841 AGTTACCGCTGGGAGATACGTCGACAGATTAATACTCAAGAGAGAGAGAGAGAGAGAG 900  
 Db 915 AGTTACCGCTGGGAGATACGTCGACAGATTAATACTCAAGAGAGAGAGAGAGAGAGAG 974  
 Qy 901 GCAAGAGCGGTGATTAATGATTAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 Db 975 GCAAGAGCGGTGATTAATGATTAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1034  
 Qy 961 GACATGACGAATTTTCTAGAGATGATTAATGATTAATGATTAATGATTAATGATTAATG 1008  
 Db 1035 GACATGACGAATTTTCTAGAGATGATTAATGATTAATGATTAATGATTAATGATTAATG 1082

```

RESULT 62
B246240/c 822 bp DNA linear GSS 13-DEC-2002
LOCUS BONJ104TR BO.1.6.2_KB_tot Brassica oleracea genomic clone BONJ104,
DEFINITION genomic survey sequence.
ACCESSION B246240
VERSION B246240.1 GI:26709961
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
  rosids; eustosids II; Brassicales; Brassicaceae; Brassica.
  1 (bases 1 to 822)
  Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
  Utzerback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
  Whole genome shotgun sequencing of Brassica oleracea and its
  application to gene discovery and annotation in Arabidopsis
  Genome Res. 15 (4), 487-495 (2005)
TITLE JOURNAL
PUBMED 15805490
COMMENT Contact: Chris Town
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA.
  Tel: 301-838-3523
  Fax: 301-838-0208
  Email: cdtown@tigr.org
  DNA is from a doubled haploid provided by Tom Osborn.
  Seq primer: TR
  Class: sheared ends.
FEATURES
  source
    1..822
    /organism="Brassica oleracea"
    /mol_type="genomic DNA"
    /strain="TO100DH3"
    /db_xref="taxon:3712"
    /clone="BONJ104"
    /note="Vector: pHD51; Site 1: BstXI; 1.6-2 kb sheared
    total DNA inserted into pHD51 using BstXI linkers"
ORIGIN
Query Match 35.2%; Score 512.8; DB 9; Length 822;
Best Local Similarity 78.1%; Pred. No. 1.2e-146;
Matches 642; Conservative 0; Mismatches 177; Indels 3; Gaps 2;
QY 455 TATCGAAGAAGAGCCGCTACACGCGCAGAGCCGCGTATACGTAGGTC 514
DB 822 TATCTGGGAAATAGCATCGCTACTCTGTGCCAGACCGCGTATATACAGAGATGC 763
QY 515 AAGGTTAATCGAATTAATCTGCGGTAAAGGGTCCGAGCGCGCTTGAAGTCAGATGT 574
DB 762 AAGGTTATCCGGAATGATTTGGGCTAAAGCGTCTAGAGTGCTTTTAAAGTCGCGGT 703
QY 575 GAAATCCCGGCTTAACTGCGAATTGCGTTGAAACTACAAAGCTAGAGTGGCAGA 634
DB 702 CAAATCCAGGGCTCAACCTCGACAGCGGTGAAACTACCAAGCTTAGAGTACGTAGG 643
QY 635 GGGAGGTGAATTCAGATGTGAGAGGAATGGGTAGATATGGAAGAACATCATGTCG 694
DB 642 GCGAGAGGAATTTCCGCTGAGCGGTAAATGGTGAAGATCGAAGAAACACCAACGG 583
QY 695 CGAAGGAGCTCTCTGGGTAAACATGACGCTCACTGACGAAGCGTGGGAGCAACAG 754
DB 582 CGAAGGAGCTCTGCTGGGCGGACATGACATGAGAGACGAAGAGTGGGAGCGAATGG 523
QY 755 GATTAGATACCTGTGATGTCACGCGCTTAAACGATGTAATTAAGTTGGGCTTATTA 814
DB 522 GATTAGATACCCCAAGTAGTCTTAGCCGTAAACGATGATCTAGGCGCTGTGCTATGCA 463
QY 815 GGCCTTG--GTAAAGAGTAAAGGTGAACCGCTGAGGAGTACGGTTCGCAAGAT 872
DB 462 CCCGTCAGTCTGTAGCTAAAGCGGTAAAGTATCCCGCTGGAGATGATCGTTCCAGAA 403

```

```

QY 873 TAAACTCAAGGAATTTGACGGGACCCGACAGCGGTGATTAATGATTAATTCA 932
DB 402 TGAACCTCAAGGAATTTGACGGGCGCCGACAGCGGTGAGCATGTGTTTAATTCGA 343
QY 933 TCGAAGCGCAAAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 992
DB 342 TCGAAGCGCAAGACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 283
QY 993 TG-CTTGGGAAACGCTTACACAGGTCTGTCATGCTGTGTCATGCTGTGTCATGCTGTG 1051
DB 282 TGCCTTCCGGAACGCGGACACAGTGTGTCATGCTGTGTCATGCTGTGTCATGCTGTG 223
QY 1052 TTGGGTTAGTCCCGAAGAGCGCAACCTTGTCTATTAATTCATTAATTCATTAATTCAT 1111
DB 222 TTGGGTTAGTCCCGAAGAGCGCAACCTTGTCTATTAATTCATTAATTCATTAATTCAT 163
QY 1112 ACTTAAATGAGACTGCGGTGACAAACCGGAGAGTGGGAGTGAAGTCAAGTCTCAT 1171
DB 162 ACCCTAAGACAGACTGCGGTGATTAACCGGAGAGTGAAGTGAAGTCAAGTCTCAT 103
QY 1172 GGCCTTAATGAGTGGGCTTACACGTAATTAATTAATTAATTAATTAATTAATTAATTA 1231
DB 102 GCCCTTAATGAGTGGGCTTACACGTAATTAATTAATTAATTAATTAATTAATTAATTA 43
QY 1232 GCGAGGGGAGGCTATCTCAAGAAAGCGCGTGTAGTCCGAT 1273
DB 42 GCGAGGGTGAAGCTATCTCAAGAAACCGGTCTTCAAGTTCGAT 1

```

RESULT 63

B2507023 824 bp DNA linear GSS 16-DEC-2002

LOCUS BONJ176TR BO.1.6.2\_KB\_tot Brassica oleracea genomic clone BONJ176,

DEFINITION genomic survey sequence.

ACCESSION B2507023

VERSION B2507023.1 GI:27028590

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eustosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 824)
 Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
 Utzerback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
 Whole genome shotgun sequencing of Brassica oleracea and its
 application to gene discovery and annotation in Arabidopsis
 Genome Res. 15 (4), 487-495 (2005)
TITLE JOURNAL
PUBMED 15805490
COMMENT Other\_GSS8: BONJ176TF
 TIGR
 Contact: Chris Town
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
FEATURES
 source
 1..824
 /organism="Brassica oleracea"
 /mol\_type="genomic DNA"
 /strain="TO100DH3"
 /db\_xref="taxon:3712"
 /clone="BONJ176"
 /note="Vector: pHD51; Site 1: BstXI; 1.6-2 kb sheared
 total DNA inserted into pHD51 using BstXI linkers"
ORIGIN
Query Match 35.2%; Score 512.8; DB 9; Length 824;
Best Local Similarity 78.1%; Pred. No. 1.2e-146;



Matches	642;	Conservative	0;	Mismatches	177;	Indels	3;	Gaps	2;
QY	484	GTGCCAGCAGCCCGGGTAATACCTAGGAGTCAGGAAACCTTAATCGAATTTACTGGGCGTAA	543						
DB	3	GTGCCAGCAGCCCGGGTAATACCTAGGAGTCAGGAAACCTTAATCGAATTTACTGGGCGTAA	62						
QY	544	GGGTGGCGCAGCGCGCTTTGAATGTCAGATGTGAATCCCGGGGCTTAACCTGGGAAATTGC	603						
DB	63	GCGTCTGTAGGTGGCTTTTAAGTCCGCCGTCAATCTCCAGGGCTCAACCTCGACAGGC	122						
QY	604	GTTTGAACCTCAAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAGTGA	663						
DB	123	GGTGAACCTCAAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAGTGA	182						
QY	664	AATGCGTAGATATGGAAGAACAATCGATGGCCGAGCGACCTCTGGGGTTAACTACTGAC	723						
DB	183	AATGCGTAGATATGGAAGAACAACAACGCGCAAAAGCACTCTGCTGGGCGGACACTGAC	242						
QY	724	GCTCATGCAAGAAAGCGTGGGGAGCAAGAAATTAGATACCTCGGTGTCACGCGCTTA	783						
DB	243	ACTGAGAGAGAAAGCTAGAGGAGCGAATGAGATAGATACCCGAGTGTCTAGCCGTA	302						
QY	784	AACGATGCACTAGTGTGTGGCCTTATTAGGCTTG--GTAAAGAACTAACGCGTGA	841						
DB	303	AACGATGCACTAGTGTGTGGCCTTATTAGGCTTG--GTAAAGAACTAACGCGTGA	362						
QY	842	GTTGACCGCTGGGAGTACGCTGCAGAAATTAAACTCAAGGAATTGACGGGACCGG	901						
DB	363	GTTATCCCGCTGGGAGTACGCTGCAGAAATTAAACTCAAGGAATTGACGGGACCGG	422						
QY	902	CACAAGCGGTGATTTATGTGATTAATTCGATGCAACGCGAAACCTTACCTACCTTG	961						
DB	423	CACAAGCGGTGATTTATGTGATTAATTCGATGCAACGCGAAACCTTACCTACCTTG	482						
QY	962	ACATGTAGCAATTTTCTAGAGATAGATTAGTG--CTTCGGGAACGCTAACACAGTGTCTG	1020						
DB	483	ACATGTAGCAATTTTCTAGAGATAGATTAGTG--CTTCGGGAACGCTAACACAGTGTCTG	542						
QY	1021	CATGCGTGTGTGACGCTGCTGTGAGATGTTGGTTAAGTCCCGCAACGACGCGAAC	1080						
DB	543	CATGCGTGTGTGACGCTGCTGTGAGATGTTGGTTAAGTCCCGCAACGACGCGAAC	602						
QY	1081	CTTGTCAATTAATTGCCATCATTTGGTTGGGCACTTTAATGACCTGCCGCTGCAAAACG	1140						
DB	603	CTTGTCAATTAATTGCCATCATTTGGTTGGGCACTTTAATGACCTGCCGCTGCAAAACG	662						
QY	1141	GAGGAAGTGGGAGTACGTCAGTCTCATGAGCCCTTATGGTGAAGGCTTCAACAGTAA	1200						
DB	663	GAGGAAGTGGGAGTACGTCAGTCTCATGAGCCCTTATGGTGAAGGCTTCAACAGTAA	722						
QY	1201	TACAAATGGCGGTACAGAGGTTGCCAACCCCGAGGGGAGCTTAATCTCAAGAAAGCGG	1260						
DB	723	TACAAATGGCGGTACAGAGGTTGCCAACCCCGAGGGGAGCTTAATCTCAAGAAAGCGG	782						
QY	1261	TGCTTAGTCGGAGTCGAGTCTGCAACTCGACTCGTGAAGTGC	1302						
DB	783	CTTCAGTTTGGATTGCAAGCTGCGACTCGCTCGATGAAAGCC	824						
RESULT 64									
BHS29859									
LOCUS	BHS29859	837 bp	DNA	linear	SSS 14-DEC-2001				
DEFINITION	BOGD080TR BOGO Brassica oleracea genomic clone BOGD080, genomic								
survey sequence.									
ACCESSION	BHS29859								
VERSION	BHS29859.1	GI:17745657							
KEYWORDS	SSS.								
SOURCE	Brassica oleracea								
ORGANISM	Brassica oleracea								
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;								
	rosids; eurosids II; Brassicales; Brassicaceae; Brassica.								
REFERENCE	1 (baee 1 to 837)								

AUTHORS	Ayele,M., Haas,B.J., Kumar,N., Mu,H., Xiao,Y., Van Aken,S., Utterback,T.R., Mortman,J.R., White,O.R. and Town,C.D. Whole genome shotgun sequencing of <i>Brassica oleracea</i> and its application to gene discovery and annotation in <i>Arabidopsis</i> Genome Res. 15 (4), 487-495 (2005)		
JOURNAL			
PUBMED	15805490		
COMMENT	Other GSSs: BOGD080TF Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TR Class: sheared ends.		
FEATURES	Location/Qualifiers		
source	1..837 /organism="Brassica oleracea" /mol_type="genomic DNA" /strain="TOL00DH3" /db_xref="taxon:3712" /clone="BOGD080" /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"		
ORIGIN			
Query Match	35.2%;	Score 512.8;	DB 9; Length 837;
Best Local Similarity	78.2%;	Pred. No. 1.2e-146;	
Matches	654;	Conservative 0;	Mismatches 177; Indels 5; Gaps 3;
QY	535 GGGCGTAAAGGTTGGCAGCGCGCTTGTAGTCAAGTGAATCCCGGCTTAACCT	594	
DB	2 GGGCGTAAAGCCTGTAGTGGCTTTTAATCCCGCTCAATCCAGGCGCTAACCC	61	
QY	595 GGGAAATGGCGTTGAACCTCAAAAGCTAGAGTGTGACAGAGGAGTGAATTCATGTG	654	
DB	62 TGGACAGCGGTGGAACTACCAAGCTTAGTACGATAGGAGCAGAGGAAATTCGCGTG	121	
QY	655 TAGCAGTGAATGCTAGATGATGAAAGAAATCATGATGCGAGCGACGCTCTGGGTT	714	
DB	122 GAGCGGTGAATGCGTGAAGATCGAAAGAACACCAACGCGGAAAGCACTGCTGGGCC	181	
QY	715 AACACTGACGCTCATGACAGAAAGCTGGGAGCAACAGATTTGATACCTGCTAGTC	774	
DB	182 GACACTGACACTGAGAGCAAGAAAGCTAGGGAGCAATGGGATTAATGATACCCACGTAGTC	241	
QY	775 CAGCGCTTAAAGATGTCAACTAGTTGTTGGGCTTTATTAGGCTTG--GTAAAGAACTA	832	
DB	242 CTAGCGCTTAAAGATGTCAACTAGGCGCTGTGCTGATGACCCGTGCACTGCTAGCTTA	301	
QY	833 ACGGTGAAGTTGACCGCGCTGGGAGTACGCTGCGAAGATTAAATCTCAAGAAATTGAC	892	
DB	302 ACGGTGAAGTATCCCGCTGGGAGTACGTTGCGAAGAAATGAATCTCAAGAAATTGAC	361	
QY	893 GGGAGCCGCAACAGCGGTGATTTATGTGATTAATTTGATGACACCGGAAAAAAGCTTAC	952	
DB	362 GGGAGCCGCAACAGCGGTGATGATGATGATTTAATTCATGCAACGAAAGAACTTAC	421	
QY	953 CTACCTTGACATGTAGGAATTTTCTAGAGATAGATTAGTG--CTTCGGGAACGCTAAC	1011	
DB	422 CAGGCTTGACATGTAGGAATTTTCTAGAGATAGATTAGTG--CTTCGGGAACGCTAAC	481	
QY	1012 CAGGTGTCATAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1071	
DB	482 CAGGTGTCATAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	541	
QY	1072 ACGCAACCTTGTCAATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1131	
DB	542 ACGCAACCTTCGTGTTTGAATGTCACGTTGAATTTGAACCTGAAACGACTGCGGCT	601	
QY	1132 GACAAACCGGAGGAGTGGGATGACGTCAAGTCTCATAGCGCTTATGGGTAGGGCTT	1191	



Db 602 GATTAACCCGAGAGAGAGTGAAGTGAAGTCAAGTCAATGATGAGCCCTTAATGCCCCTGGCCGA 661  
Qy 1192 CACACGTAATTAACAATGGCGGTACAGAGGGTTGCCAACCCGCGAGGGGGAGTAAATCTCA 1251  
Db 662 CACACGTGTACATAGCCCGGACAAAGGTCCGATCCCGCAGAGGTAGTAACTCCA 721  
Qy 1252 GAAGCGCGTCTGATGTCGGATCGAGTGCACACTGCACTCCGTGAAGTGGAAATCGCT 1311  
Db 722 AAAACCGCTCTCAATTCGATTCGAGGTGCAATCTGCTGCATGAAGCCGGAATCGCT 781  
Qy 1312 AGTAATCGCGGATCAGCATGTC--GGGTAAATACGTTCCCGGCTCTTGAACAC 1365  
Db 782 AGTAATCGCGGTCAAGCATACGCGGTGAATTCGTTCCCGGCTTGAACAC 837

RESULT 65  
B2435283 862 bp DNA linear GSS 13-DEC-2002  
LOCUS BONHF62TR BQ.1.6.2 KB tot Brassica oleracea genomic clone BONHF62.  
DEFINITION genomic survey sequence.  
ACCESSION B2435283  
VERSION B2435283.1 GI:26686341  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 862)  
Ayle, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Utecherback, T.R., Mortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
JOURNAL PUBMED 15805490  
COMMENT Other GSSs: BONHF62TR  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.

FEATURES  
Location/Qualifiers  
source 1..862  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="T01000DH3"  
/db\_xref="taxon:3712"  
/clone="BONHF62"  
/clone\_id="BO\_1.6.2 KB tot"  
/note="Vector: PHOS1; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into PHOS1 using BstXI linkers"

ORIGIN  
Query Match 35.2%; Score 512.8; DB 9; Length 862;  
Best Local Similarity 78.9%; Pred. No. 1.2e-146;  
Matches 636; Conservative 0; Mismatches 167; Indels 3; Gaps 2;

Qy 624 AGTGTGGCAGAGGAGAGTGGAAATTCATGTGTAGAGTGAATGCGTAGATATGAG 683  
Db 236 AGTACGTTGTGGGGCAGAGAGGAATTTTCGGTGAAGCGGTAAATGCTTAAGATCGAAG 295  
Qy 684 AACATCATGTGGGAGGAGGAGCCTCTCGGTAACTGAAGCTGATGACGAAAGCGTG 743  
Db 296 AACATCAAGCGGAGGAGGAGCCTCTCGGTAACTGAAGCTGATGACGAAAGCGTG 355  
Qy 744 GGAGCAAGAGATTAATGATACCTCGGTAGTCCAGCGCTTAAACGATGTCAATGTTT 803  
Db 356 GGAGCAAGATGAGATTAATGATACCTCGGTAGTCCAGCGCTTAAACGATGTCAATG 415  
Qy 804 GGGCTTATTAAGCTTG--GTAAAGAGTAAAGCGGTAAAGTTGACGCGCTGGAGATAC 861  
Db 416 GTCCGTATCGACCCCGTCAAGTGTGATGCTAAACGCGTAAAGTATCCCGCTGGAGATAC 475  
Qy 862 GGTGCGAAGATTAAACTCAAGAGATTAAGCGGGGACCCGCAACAGCGGTGATATG 921  
Db 476 GTTCGCAAGATTAATCAAGAGATTAAGCGGGGACCCGCAACAGCGGTGATATG 535  
Qy 922 GATTAAATTCGATGCAACGCGGAAACCTTACTTACCTTGAATGAGCAATTTCTAG 981  
Db 536 GTTAAATTCGATGCAACGCGGAAACCTTACTTACCTTGAATGAGCAATTTCTAG 595  
Qy 982 AGATGATTTATGTC--CTTCGGAAACGCTAACACAGGTGCTGCAATGCTGTCAGCTCGT 1040  
Db 596 AAGAGAGAGGGGTGCTCTTCGGGAAACGCGGACACAGGTGTCATGCTGCTGTCAGCTCGT 655  
Qy 1041 GTGCTGATATGTTGGGTAAAGTCCGCAACGAGCGCAACCTTGTCAATTAATGGCATCA 1100  
Db 656 GCGGTAAAGGTGTGGGTAAAGTCCGCAACGAGCGCAACCTTGTCAATTAATGGCATCA 715  
Qy 1101 TTGTGTGGGCACTTTAATGAGACTGCGGTGACAAACCGGAGAAAGTGGGATGACGT 1160  
Db 716 TTGAGTTTGGACCTTGAAACGAGACTGCGGTGATTAAGCCGAGAAAGTGAAGATGACGT 775  
Qy 1161 CAAGTCTCTATGCGCCCTTATGAGGTGAGGCTTCAACGATTAATGAGCGCGATACAGAG 1220  
Db 776 CAAGTCAATCAATGCGCCCTTATGAGGTGAGGCTTCAACGATTAATGAGCGCGATACAGAG 835  
Qy 1221 GTTGCCAAACCGGAGGAGGAGCTAA 1246  
Db 836 GTGCGATCCCGAGGAGTAACTAA 861

RESULT 66  
BH582045/c 861 bp DNA linear GSS 15-DEC-2001  
LOCUS BOGZU46TR BOGZ Brassica oleracea genomic clone BOGZU46, genomic  
DEFINITION survey sequence.  
ACCESSION BH582045  
VERSION BH582045.1 GI:17834502  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 861)  
Ayle, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Utecherback, T.R., Mortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
JOURNAL PUBMED 15805490  
COMMENT Other GSSs: BOGZU46TR  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF  
Classes: sheared ends.

# FEATURES

source

```
1..861
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOG2046"
/clone_1b="BOG2"
/notes="Vector: PHOS1, Site 1: BexXI; 2-3 kb sheared
genomic DNA inserted into phos1 using BexXI linkers"
```

## ORIGIN

Query Match 35.2%; Score 512.4; DB 9; Length 861;  
Best Local Similarity 76.6%; Pred. No. 1.7e-146;  
Matches 677; Conservative 0; Mismatches 181; Indels 26; Gaps 3;

```
243 AGGCCAGCATCAGTACTGTGTGAGAGACGACCGACCTGGAGCTGAGACACGGC 302
Db 861 AGCCGATGATCAGTACTGTGTGAGAGACGACCGACCTGGAGCTGAGACACGGC 802
Qy 303 CCAGACTCTTACGCGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362
Db 801 CCAGACTCTTACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 742
Qy 363 GCATGCGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 422
Db 741 GCATGCGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 684
Qy 423 AGGTTACGTTAATATCGGACCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 482
Db 683 -----ACGATGACGGATATCTGGGAAATAGCATCGCTTAACTC 645
Qy 483 CGTCCAGAGCGCGCGTAAATAGTGGGTCAGAGCGTTAATCGAATTACTGGCGCTAA 542
Db 644 TGTGCGAGAGCGCGCGTAAATAGTGGGTCAGAGCGTTAATCGAATTACTGGCGCTAA 585
Qy 543 AGGTCGCGAGCGCGCTTAAATAGTGGGTCAGAGCGTTAATCGAATTACTGGCGCTAA 602
Db 584 AGCGTCTAGGAGCGCTTAAATAGTGGGTCAGAGCGTTAATCGAATTACTGGCGCTAA 525
Qy 603 CGTTTGAATCTCAAGCTTAAAGTGGGTCAGAGCGTTAATCGAATTACTGGCGCTAA 662
Db 524 CGGTGAAATCTCAAGCTTAAAGTGGGTCAGAGCGTTAATCGAATTACTGGCGCTAA 465
Qy 663 AATGCGTAAAGTATGAAAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
Db 464 AATGCGTAAAGTATGAAAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 405
Qy 723 CGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
Db 404 CACTGAGAGACGAAAGCTAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 345
Qy 783 AAAAGATGCACTAATGTTGGGCTTATTAAGCTTG--GTAAAGAGTAAAGGAGGAG 840
Db 344 AAAAGATGCACTAATGTTGGGCTTATTAAGCTTG--GTAAAGAGTAAAGGAGGAG 285
Qy 841 AGTTAGCCGCTGGGAGGAGTACGGTCCGAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 900
Db 284 AGTATCCGCTGGGAGGAGTACGGTCCGAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 225
Qy 901 GCACAAAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 960
Db 224 GCACAAAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 165
Qy 961 GACATGTACGGAATTTTCTAAGATTAATGATTAATGATTAATGATTAATGATTAATG 1019
Db 164 GACATGTACGGAATTTTCTAAGATTAATGATTAATGATTAATGATTAATGATTAATG 105
Qy 1020 GATGAGCTGCTGACGCTGCTGAGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1079
```

Db 104 GCATGCGTGTCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 45  
Qy 1080 CCTGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1123  
Db 44 CCTGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1

RESULT 67  
DR027023 895 bp mRNA linear EST 26-MAY-2005  
LOCUS Osmo01058 F. cylindrus osmotic stress library Fragilariopsis  
DEFINITION -cylindrus cDNA clone FcylESTa47e01.s1, mRNA sequence.  
ACCESSION DR027023  
VERSION DR027023.1 GI:66749386  
KEYWORDS EST  
SOURCE Fragilariopsis cylindrus  
ORGANISM Fragilariopsis cylindrus  
Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;  
Bacillariophycidae; Bacillariales; Bacillariaceae; Fragilariopsis.

REFERENCE  
AUTHORS Krell, A. and Gloeckner, G.  
TITLE Analysis of an osmotic stress induced cDNA library of the  
psychrophilic diatom Fragilariopsis cylindrus  
JOURNAL Unpublished (2004)  
COMMENT Contact: Krell, Andreas; Gloeckner, Gernot  
Biological Oceanography, Sea ice research; Genome Analysis  
Alfred-Wegener-Institute for Polar and Marine Research; Institute  
for Molecular Biotechnology  
Am Handelshafen 12, D-27570 Bremerhaven, Germany; Beutenbergstr.  
11, D-07745 Jena, Germany  
Tel: ++49 471 48311812; ++49 3641 656440  
Fax: ++49 471 48311425; ++49 3641 656255  
Email: akrell@awi-bremerhaven.de; gernot@imb-jena.de

PCR Primers  
FORWARD: 5'M13  
BACKWARD: 3'M13  
Seq primer: 5'GTAAACGACGCGCAG 3'.  
Location/Qualifiers

## FEATURES

source

```
1..895
/organism="Fragilariopsis cylindrus"
/mol_type="mRNA"
/db_xref="taxon:186039"
/clone="FcylESTa47e01.s1"
/clone_1b="F. cylindrus osmotic stress library"
/notes="Samples for total RNA isolation were taken  
continuous for 5 days after a salt shock treatment  
increasing salinity from 34 to 60 PSU. Total RNA  
extraction was performed with RNeasy (Ambion) and mRNA  
purification with polyA Purist (Ambion). Further steps  
were carried out as described in the Clontecher Kit. cDNA  
size fractionation was carried out with CHROMA Spin-400  
columns and additionally on a gel."
```

## ORIGIN

Query Match 35.1%; Score 511.6; DB 8; Length 895;  
Best Local Similarity 75.4%; Pred. No. 3e-146;  
Matches 683; Conservative 0; Mismatches 198; Indels 25; Gaps 3;

```
196 GAGCGCGCGATGCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 255
Db 2 GAGAGGCTTGGCGGCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 61
Qy 256 TAGTGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 315
Db 62 TAGTGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121
Qy 316 GAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 375
Db 122 GAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 181
Qy 376 AGTGAAGAGGCGCTTGGGTTAAAGCTCTTTCACTGAGGAGGAGGAGGAGGAGGAG 435
Db 182 AGGATGACGCTATGAGGTTAAAGCTCTTTTCAAGGAGGAGGAGGAGGAGGAGGAGG 236
```

QY 436 TAATGTGATCCCATGACGGTATGACAGAGAGACCGGCTTAAGTCTGCTCCAGACGCC 495  
 DB 237 -----TGTACCTGAGATTAAGCATCGGCTTAAGTCTGCTCCAGACGCC 279  
 QY 486 GCGGTAAATCGTAGGGTGTCAACCGTTAATCGGAATTAAGTGGCGTAAAGGGCGGACGCC 555  
 DB 280 GCGGTAAATCGTAGGGTGTCAACCGTTAATCGGAATTAAGTGGCGTAAAGGGCGTAAAGGGT 339  
 QY 556 GGCCTTGAATGACAGATGGAATCCCGGGCTTAAGTCTGGAATTCGCTTGAATTAAGTCT 615  
 DB 340 TGTGTAATGATGCACTGTTAATCTTGAAGTCACTTAAGTCACTTAAGTCACTTAAGTCACT 399  
 QY 616 AAGGTAGAGTGTGACAG 675  
 DB 400 CAGACTAGAGTAAATGAG 459  
 QY 676 TATGAG 735  
 DB 460 TTGAG 519  
 QY 736 AAGCGTGGGAG 795  
 DB 520 AAGCTAG 579  
 QY 796 TAGTGTGGGCTTATTAAGGCTTG--GTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 853  
 DB 580 TAGATGTTAAG 639  
 QY 854 GGGAGTACGGTGTGAG 913  
 DB 640 GGAAGTATCTCCAG 699  
 QY 914 ATTATGTGATTAATCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 973  
 DB 700 AGCATGTGTTAATGATGAG 759  
 QY 974 TTTCTAGAGATGATTAAGTGTG--CTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1032  
 DB 760 TTTCTTGAAG 819  
 QY 1033 CAGCTGTGTGTGAG 1092  
 DB 820 CAGCTGTGTGTGAG 879  
 QY 1093 TGGCAT 1098  
 DB 880 TGCCTT 885  
 RESULT 68  
 BH484054 819 bp DNA linear GSS 13-DEC-2001  
 LOCUS BOHGA75TF BOHG Brassica oleracea genomic clone BOHGA75, genomic  
 DEFINITION survey sequence.  
 ACCESSION BH484054  
 VERSION BH484054.1 GI:17692158  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Brassicaceae; Streptophyta; Embryophyta; Tracheophyta;  
 Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 819)  
 Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
 Utechtack, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 JOURNAL PUBMED 15805490  
 COMMENT Other\_GSSs: BOHGA75TF  
 Contact: Chris Town  
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.  
 Location/Qualifiers  
 source 1..819  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="T01000DH3"  
 /db\_xref="taxon:3712"  
 /clone\_lib="BOHGA75"  
 /clone="BOHG"  
 /note="vector: pHOSt, Site 1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHOSt using BstXI linkers"  
 ORIGIN  
 Query Match 35.1%; Score 511.4; DB 9; Length 819;  
 Best Local Similarity 78.1%; Pred. No. 3.3e-146;  
 Matches 640; Conservative 0; Mismatches 176; Indels 3; Gaps 2;  
 QY 512 TCGAAGCGTTAATCGGAATTAATCTGAGCGGTAAAGGGTGCAGCGCGCTTGAATGACAG 571  
 DB 1 TCGAAGCGTTAATCGGAATTAATCTGAGCGGTAAAGGGTGCAGCGCGCTTGAATGACAG 60  
 QY 572 TGTGAATCCCGGGCTTAACTGGGAATTCGTTTGAATCAATCAAGTCAAGTGTGCG 631  
 DB 61 CGTCAATCCAGGGCTCAACCTTGACAGGGGTGGAATCAACAGCTTGAATGACGT 120  
 QY 632 AAGAGAGGTGAATTCATGTGTAGCAGTGAATGCGTAGATATGAGAGATGAGAGACATCA 691  
 DB 121 AAGGGGAGAGGAAATTTCCGTGAGAGCGGTGAATGCGTAGATATGAGAGATGAGAGAC 180  
 QY 692 TGGCGAGGACAGCTCTCGGTTAACTAGAGCTCATGACAGAAAGCGTGGGAGACAA 751  
 DB 181 CCGGAGAGACATCTGCTGGGCGGACACTGACATGAGAGAGAAAGTAAAGGAGACGAA 240  
 QY 752 CAGATTAATGATCCCTGTGATGACAGCGCTTAAAGATGATCAATGATGTTGGCTTA 811  
 DB 241 TGGGATTAATGATCCCTGTGATGACAGCGCTTAAAGATGATCAATGATGTTGGCTTA 300  
 QY 812 TTAGGCTTG--GTAAAGAGCTTAACCGGTGAAGTGAACCGCTGGGAGATGAGTGCAG 869  
 DB 301 CGACCGGTGAGGTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 QY 870 GATTAATCAAGAGATGAG 929  
 DB 361 GATTAATCAAGAGATGAG 420  
 QY 930 CGATGCAAGCGGAG 989  
 DB 421 CGATGCAAGCGGAG 480  
 QY 990 TAGTG--CTTGGAGAGCTTAACAGAGTGTGATGAGTGTGCTGAGTGTGCTGAG 1048  
 DB 481 GGGTGCCTTCGGAGAGCGGAGACAGAGTGTGATGAGTGTGCTGAGTGTGCTGAG 540  
 QY 1049 ATGTTGGTTAATCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1108  
 DB 541 GTTGGTTAATCCCGGAG 600  
 QY 1109 GGCATTTAATGAGACTGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1168  
 DB 601 GGAACCTGAAG 660  
 QY 1169 CATGCGCTTATGAG 1228  
 DB 661 CATGCGCTTATGAG 720  
 QY 1229 CCGGAGAGGGGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1288

Db 721 CCCGAGGGGTGAGCTAACTCCAAAAACCCGTCCTCAGTTCGAGATTGACGCTGCACATC 780

Qy 1289 GACTCGTGAAGTCGGAATCGCTAGTAATCGCGAGATCAG 1327

Db 781 GCTGTGATGAAGCCGGAATCGCTAGTAATCGCGGCTCAG 819

RESULT 69

LOCUS BZ468986 935 bp DNA linear GSS 13-DEC-2002

DEFINITION BONKD78TF BO\_1.6\_2\_KB\_tot Brassica oleracea genomic clone BONKD78, genomic survey sequence.

ACCESSION BZ468986

VERSION BZ468986.1 GI:26764518

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE 1 (bases 1 to 935) Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utecherback, T.R., Wortman, J.R., White, O.R. and Town, C.D. Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res. 15 (4), 487-495 (2005)

JOURNAL PUBMED 15805490

COMMENT Other\_GSSs: BONKD78TF

CONTACT: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.

TEL: 301-838-3523

FAX: 301-838-0208

EMAIL: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

FEATURES

source location/Qualifiers

1..935

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="TO1000DH3"

/db\_xref="taxon:3712"

/clone="BONKD78"

/clone\_lib="BO\_1.6\_2\_KB\_tot"

/note="Vector: PHOSI; Site\_1: BstXI; 1.6-2 kb sheared total DNA inserted into PHOSI using BstXI linkers"

ORIGIN

Query Match 35.1%; Score 511; DB 9; Length 935;

Beet Local Similarity 77.6%; Pred. No. 4,6e-146; Indels 4; Gaps 3;

Matches 656; Conservative 0; Mismatches 185;

Qy 617 AAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGAGTGAATGCGTAGAGAT 676

Db 1 AAGCTTAGTAGTAGAGGCGAGAGGGAATTCGCGTGAAGCGTGAATGCGTAGAGAT 60

Qy 677 ATGGAAGACATGATGCGGAAGGAGCCTCTGCGTTAACTGACGTACGTCACGAA 736

Db 61 CGGAAGAACACCAACGCGGAAGGAGCCTCTGCGCGACACTGACACTGAGAGACGAA 120

Qy 737 AGCGTGGGAGGAGAACAGATTAAGTACCTCTGTGTGTCCAGCCCTTAACGATGTCACT 796

Db 121 AGCTTAGGGGAGGAGTGGATTAGATACCCAGTAGTCTTACCGGTAAACGATGATCT 180

Qy 797 AGTTGTGGGCTTATTAGGCTTG--GTAACGAGACTAACGCGTGAAGTTGACCGCTG 854

Db 181 AGGCGCTGTGCTATCGACCCGTCGACGTCTGTAGCTAACGCGTTAAGTATCCGCGCTG 240

Qy 855 GGAGTACGCTGCGCAAGATTAAACTCAAGGAATTGACGCGGACCCGCAACGCGGTGA 914

Db 241 GGAGTACGCTGCGCAAGATTAAACTCAAGGAATTGACGCGGCGCCGCAACGCGGTGA 300

Qy 915 TTATGTGATTAAATTCGATGCAACGCGAAAAACCTTACCTTACCTTGAATGACGAAT 974

Db 301 GCATGTGTTTAATTCGATGCAACGCGAAAAACCTTACCTTACCTTGAATGACGAAT 360

Qy 975 TTTCTAGAGATAGATTAGTG-CTTCGGGAGAGCTTAACACAGAGTGTGATAGGCTGCTC 1033

Db 361 CCTCTTGAAGAGAGGGGTGCTTCGGGAGAGCGGACACAGTGTGATAGGCTGCTGCTC 420

Qy 1034 AGCTGTGTCGTGAGATGTTGGTTAAGTCCCGCAACGAGGCGCAACCTTGTCAATTT 1093

Db 421 AGCTGTGTCGTGAGATGTTGGTTAAGTCCCGCAACGAGGCGCAACCTTGTGATTT 480

Qy 1094 GCCATCATTTTGTGTTGGACCTTTAATGAGACTGCCGCTGACAAACCGGAGGAAGTGGG 1153

Db 481 GCCACCGTTGAGTTTGAACCCCTGAAACGACTGCCGCTGATTAAGCCGAGGAAGGTAGG 540

Qy 1154 ATGACGTAAATGCTTCATAGCCCTTATAGGTTAGGGCTTCAACGTTAATCAATGGCGGT 1213

Db 541 ATGACGTCAATGCTTCATAGCCCTTATAGCCCTTGGGCGACACAGCTGCTAATAGCCCGG 600

Qy 1214 AAGAGGGTTGCGCAACCGGAGGGGAGCTAATCTGAGAAAGCGCTGTAGTCCGGAT 1273

Db 601 ACAAAGGTTGCGCAATCCGCGAGGGTGAAGCTTAACCTCAAAAACCGCTCTTCAATTCGAT 660

Qy 1274 CGSAGTCTGCAACTGCACTCCGTAAGTCGGAATCGCTAGTAATCGCGGATCAG-CATGT 1332

Db 661 TGCAGGCTGCAACTGCTGCTGATGAAGCGGAATCGCTAGTAATCGCGGATCAGCATAC 720

Qy 1333 CGCGGTGAATAGCTTCCCGGCTTTGTACACACCGCCCGTCAACATGGAGTGGGTTT 1392

Db 721 GCGCGTGAATAGCTTCCCGGCTTTGTACACACCGCCCGTCAACATGGAGTGGGCA 780

Qy 1393 CACCAAGACGAGTAGTACTACCGTAAGGAGGGGCTTCCACGCTGATTAATGACTG 1452

Db 781 TGCCGGAAGTGTACTTAACCGCAAGAGGGGGTGCAGAAAGCAGGCTAGTACTG 840

Qy 1453 GGGTG 1457

Db 841 GAGTG 845

RESULT 70

LOCUS BH453601/c 865 bp DNA linear GSS 12-DEC-2001

DEFINITION BOHBW70TF BOHB Brassica oleracea genomic clone BOHBW70, genomic survey sequence.

ACCESSION BH453601

VERSION BH453601.1 GI:17639312

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE 1 (bases 1 to 865) Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utecherback, T.R., Wortman, J.R., White, O.R. and Town, C.D. Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res. 15 (4), 487-495 (2005)

JOURNAL PUBMED 15805490

COMMENT Other\_GSSs: BOHBW70TF

CONTACT: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.

TEL: 301-838-3523

FAX: 301-838-0208

EMAIL: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

FEATURES

source location/Qualifiers

1..865

```
/organism="Brassica oleracea"  
/mol_type="genomic DNA"  
/strain="T0100DH3"  
/db_xref="taxon:3712"  
/clone="BOHBW70"  
/clone_1ib="BOHB"  
/note="Vector: pHOSt; Site_1: BstXI; 2-3 kb sheared  
genomic DNA inserted into pHOSt using BstXI linkers"
```

## ORIGIN

```
Query Match      35.0%; Score 510.2; DB 9; Length 865;  
Best Local Similarity 76.9%; Pred. No. 8e-146;  
Matches 684; Conservative 0; Mismatches 178; Indels 27; Gaps 4;
```

```
QY 279 GCCACACTGGGACTGAGACACGGCCCACTCTTACCGGAGCAGCAGTGGGGAATTTG 338  
DB |||||||  
DB 865 GCCACACTGGGACTGAGACACGGCCCACTCTTACCGGAGGAGGAGCAGT-GGGGAATTTTC 807  
QY 339 GACAAATGGGCGCAAGCTTGATCCAGCAATGCCGTGATGAGAAAGGCTTCGGGTGT 398  
DB |||||||  
DB 806 CGCAATGGGCGCAAGCTTGATCCAGCAATGCCGTGATGAGAAAGGCTTCAGGGTCT 747  
QY 399 AAGCTCTTTCAGTCGAGAGAAAGGTTACGGTAATAATCGTGAACCATGACGGTATC 458  
DB |||||||  
DB 746 GAACTTCTTTCCAGAGAGA-----AGCAATGACGGTATC 710  
QY 459 GACAGAGAGACACCGGCTTAATCTGTCGACGACCGCGGTAAATACGTAGGGTGCAAGC 518  
DB |||||||  
DB 709 TGGGGAAATAGCAATCGGCTAATCTGTGCCAGACGCGCGTAAATACAGAGATGCAAGC 650  
QY 519 GTTAATCGAATTACTGGGCGTAAAGGGTGGCGACGCGCTTGTAACTCAGATGTGAA 578  
DB |||||||  
DB 649 GTTATCCGGAATGATGGGCGTAAAGCGTGTAGTGGTCTTTTAAGTCCGCGTCAA 590  
QY 579 TCCCGGCGCTTAACCTGGGAATGGCTTGAATACTAAGCTAGATGGCGAGAGGA 638  
DB |||||||  
DB 589 TCCAGGGCTCAACCTGAGACAGCGGTGGAACTACCAAGCTTGAAGACGTAGGGCA 530  
QY 639 GGTGAATTCATGTGTAGACAGTGAATGCTAGATATGAGAAACATCGATGGCGAA 698  
DB |||||||  
DB 529 GAGGGAATTTCCGGTGAAGGGTGAATGGGTAGATCGAAAGAACACCAACGGCGAA 470  
QY 699 GGCAGCTCTCTGGGTTAACTGACGCTCATGCAAGAAAGCGTGGGAGCAACAGATT 758  
DB |||||||  
DB 469 AGCACTCTGCTGGGCGCACACTGACATGAGAGAGAAAGCTAGGGAGCGCAATGGGATT 410  
QY 759 AGATACCTCTGCTGAGTCCAGCGCTTAAAGATGTCACTAGTGTGGGCTTATTAGGCT 818  
DB |||||||  
DB 409 AGATACCTCTGAGTCTAGCCGTAAACGATGATCTAGGCGCTGTGCTATGACCCG 350  
QY 819 TG--GTAAAGAGCTAACGCGTGAAGTTGACCGCCTGGGAGTACGGTCCGCAAGATTAA 876  
DB |||||||  
DB 349 TGCAGTCTGTAGTTAACGCTTAAAGTATCCCGCTGGGAGTACGTTCCGCAAGATGAA 290  
QY 877 ACTCAAAGGAATTGACGGGAGCCCGCAAGAGCGGTGAATTATGTGATTAATTCGATGCA 936  
DB |||||||  
DB 289 ACTCAAAGGAATTGACGGGAGCCCGCAAGAGCGGTGAATGTGTTAATTCGATGCA 230  
QY 937 ACGGAAAAAATCTTACTACCTTGAATGATGAGGAATTTCTTAGAGATAGATTAGTG-C 995  
DB |||||||  
DB 229 AAGGGAAGAACCTTACAGAGGCTTGAATGACGCGGAATCTCTTGAAGAGAGGGGTGCC 170  
QY 995 TTGGGAAGCTAACAGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055  
DB |||||||  
DB 169 TTCCGGAAGCGGAGACAGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 110  
QY 1056 GTTAAGTCCCGCAACGAGCGCAACCTTGTCAATTAATTCATCAATTTGTTGGCACTT 1115  
DB |||||||  
DB 109 GTTAAGTCCCGCAACGAGCGCAACCTTGTCAATTAATTCATCAATTTGTTGGCACTT 50  
QY 1116 TAAATGAGCTGCGGTGACAAACCGGAGGAAGTGGGGATGACGTCAAG 1164  
DB |||||||  
DB 49 TGAACAGACTGCGGTGATTAAGCCGAGGAAGATGAGATGACGTCAAG 1
```

Search completed: May 5, 2006, 17:00:21  
Job time : 6232 secs

**THIS PAGE LEFT BLANK**



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/642,229A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/246,865  
FILING DATE: 20-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sheiness, Diana K.  
REGISTRATION NUMBER: 35,356  
REFERENCE/DOCKET NUMBER: UOFW19233  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 682 8100  
TELEFAX: (206) 224 0779  
TELEX: 4938023  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1536 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
DESCRIPTION: "16S ribosomal DNA"  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Comamonas testosteroni ATCC No. 5874291 11996  
US-08-642-229A-1

Query Match 77.2%; Score 1124.4; DB 2; Length 1536;

Best Local Similarly 87.1%; Pred. No. 0;

Matches 1271; Conservative 0; Mismatches 181; Indels 8; Gaps 3;

QY 1 ATTTGAAGCTGGGCGGCGATGCTTTACACATGCAAGTCGAAAGCGGACGCGATGCTTGCAT 60  
DB 29 ATTGAAGCTGGGCGGCGATGCTTTACACATGCAAGTCGAAAGCGGACGCGATGCTTGCAT 85  
QY 61 CTGGTGGCGAGTGGCGGACGGGTGATGATGATCGGAACGATTCGAAAGCGGCGGTA 120  
DB 86 -TGCTGACGATGGCGGACGGGTGATGATGATCGGAACGATTCGAAAGCGGCGGTA 144  
QY 121 ACCGATCGAAAGATGCTTAATACCGCATATCTCTAAGAGGAAAGCAGGCGATCGAA 180  
DB 145 ACTACTCGAAAGATGCTTAATACCGCATATCTAAGAGGAAAGCAGGCGATCGAA 204  
QY 181 GACCTGGCGCTTTGGAGCGGCGGATGCTGATTAGCTAGTGGGTGGGTTAAAGGCTTAC 240  
DB 205 GGCCTGTGCTACTAGAGCGGCGGATGCTGATTAGCTAGTGGGTGGGTTAAAGGCTTAC 264  
QY 241 CAAGCGACGATGATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
DB 265 CAAGCGGCGATGCTTAATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 324  
QY 301 GCCCAGACTCTTAACGGAG 360  
DB 325 GCCCAGACTCTTAACGGAG 384  
QY 361 CAGCAATGCGGCGGATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
DB 385 CAGCAATGCGGCGGATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 444  
QY 421 AAAGTTACGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
DB 445 AAAGTTACGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 504  
QY 481 TACGTGCGACAGCGCGGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
DB 505 TACGTGCGACAGCGCGGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 564  
QY 541 AAAGGTCGCGAGCGGCGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 600

DB 565 AAAGGTCGCGAGCGGCTTGTGAAGACAGTGTGAATCCCGGCGCTCAACCTGGGAGAC 624  
QY 601 TCGGTTGAATCTCAAAAGCTAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
DB 625 TCGGTTGAATCTCAAAAGCTAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 684  
QY 661 TGAATGCTAGAGATATGAG 720  
DB 685 TGAATGCTAGAGATATGAG 744  
QY 721 GACGCTCATGACGAAAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
DB 745 GACGCTCATGACGAAAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804  
QY 781 CTAAACGATGATCACTAGTTGTTGGGCTTATTAAGCTTGTGAAGAGAGAGAGAGAGAG 840  
DB 805 CTAAACGATGATCACTAGTTGTTGGGCTTATTAAGCTTGTGAAGAGAGAGAGAGAGAG 864  
QY 841 AGTTGACCGCTGGGAG 900  
DB 865 AGTTGACCGCTGGGAG 924  
QY 901 GCACAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 925 GCACAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 984  
QY 961 GACATGTCGAGAAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1017  
DB 985 GACATGTCGAGAAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1044  
QY 1018 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
DB 1045 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104  
QY 1105 ACCCTGTGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1163  
DB 1138 CCGAG 1197  
QY 1164 CCGAG 1223  
DB 1198 TAAATCAATGAGCGGATGAG 1257  
QY 1224 TCAATCAATGAGCGGATGAG 1283  
DB 1258 GCGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317  
QY 1284 CAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343  
DB 1318 CCGGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1377  
QY 1344 CCGGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1403  
DB 1378 CAGGAG 1437  
QY 1404 CAGGAG 1463  
DB 1438 TGAGATTCATGACTGGGCTG 1457  
QY 1464 CCGGCTTGTGATGCTGGGCTG 1483

RESULT 2  
US-09-726-774-5  
; Sequence 5, Application US/09726774  
; Patent No. 6677153  
; GENERAL INFORMATION:  
; APPLICANT: Iversen, Patrick L.  
; TITLE OF INVENTION: Antisense Antibacterial Method and  
; TITLE OF INVENTION: Composition  
; FILE REFERENCE: 0450-0032.30  
; CURRENT APPLICATION NUMBER: US/09/726,774



;; CURRENT FILING DATE: 2000-11-29  
;; PRIOR APPLICATION NUMBER: US 60/168,150  
;; PRIOR FILING DATE: 1999-11-29  
;; NUMBER OF SEQ ID NOS: 139  
;; SOFTWARE: PasteSeq for Windows Version 4.0  
;; SEQ ID NO: 5  
;; LENGTH: 1544  
;; TYPE: DNA  
;; ORGANISM: Neisseria gonorrhoea  
US-09-726-774-5

Query Match 76.7%; Score 1118.2; DB 3; Length 1544;  
Best Local Similarity 87.1%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 183; Indels 6; Gaps 4;

```
QY 1 ATTGAAGCTGGGGGCGGCTTTTACATGCAAGTCGACGCGACAC--GGATGCTTGC 58
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 29 ATTGAAGCTGGGGGCGGCTTTTACATGCAAGTCGACGCGACACGCGACAGGAAAGCTTGC 88
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59 ATCT--GATGCGAGTGGCGGACCGGGTGAATGATCATCGGACGTAATCCAGAAAGGGG 116
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 89 TTCTCGGGTGGCGAGTGGCGGACCGGGTGAATGATCATCGGACGTAATCCAGGAGCTTGC 148
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 117 GGTAAAGCATGAAAGATGCTTAATACCGCATTAATCTTAAGAGGAAAGCAGGGGATC 176
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 149 GATAACTGATGAAAGATCAGCTAATACCGCATACGCTTGAAGAGGGAAACAGGGGAC 208
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 177 GAAAGACCTTGGCGCTTTTGAAGCGGCGGATGCTGAATTAAGTATGGTGGGTTAAAGGC 236
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 209 TTCGGGCTTGGCGCTTTCGATCCGACGCGCATATCTGATTAAGCTGGTGGGTTAAAGGC 268
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 237 CTACCAAGGCGACGATCAGTATGTTGCTGAGAGGACGACCGACACTGGGACTGAGA 296
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 269 CCACCAAGGCGACGATCAGTATGTTGCTGAGAGGATGATCCGACACTGGGACTGAGA 328
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 297 CACGCGCCGACACTCTTACCGGAGCGACAGTGGGGAATTTTGAAGTATGGGCGCAAGCT 356
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 329 CACGCGCCGACACTCTTACCGGAGCGACAGTGGGGAATTTTGAAGTATGGGCGCAAGCT 388
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 357 GATCAGCAATGCGCGGCGGATGAAAGAGGCTTCGGGTTGAAGCTCTTTCAGTCCAG 416
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 389 GATCAGCAATGCGCGGCGGATGAAAGAGGCTTCGGGTTGAAGCTCTTTCAGTCCAG 448
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 417 AAGAAAAGGTTACGGTAAATATATCGTGAACCATGACGCTATCGACAGAAAGACACCGGC 476
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 449 AAGAAAAGGTTACGGTAAATATATCGTGAACCATGACGCTATCGTGAAGAAATAGACCGGC 508
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 477 TAACCTAGCTGCACGACGCGCGGTAAATAGTGGTGCAGCGTTAATCGAATTAATCGG 536
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 509 TAACCTAGCTGCACGACGCGCGGTAAATAGTGGTGCAGCGTTAATCGAATTAATCGG 568
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 537 GCGTAAAGGTTGCGGACGCGCTTGTAAATGATGAAATGCGGCGGCTTAATCGG 596
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 569 GCGTAAAGGTTGCGGACGCGCTTGTAAATGATGAAATGCGGCGGCTTAATCGG 628
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 597 GAATGCGCTTGAACCTAACAAGCTAGAGTGTGAGAGAGGAGGAGTGAATTCATGTTGA 656
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 629 GAATGCGCTTGAACCTAACAAGCTAGAGTGTGAGAGAGGAGGAGTGAATTCATGTTGA 688
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 657 GCAGTGAATGCGTGAAGATATGAGAAACATGATGCGGAGGAGGAGGAGGAGGAGGAGG 716
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 689 GCAGTGAATGCGTGAAGATATGAGAAACATGATGCGGAGGAGGAGGAGGAGGAGGAGG 748
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 717 CACTGACGCTATGACCAAGAACGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 776
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 749 CACTGACGCTATGACCAAGAACGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 808
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 777 GCGCCCTAAACGATGATCACTAGTGTGGGCTTATTA--GGCTTGGTAAAGAGCTAAG 835
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 809 GCGCCCTAAACGATGATCACTAGTGTGGGCAACTGATGATGCTTGGTAAAGAGCTAAG 868
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 836 CGTGAAGTTGACCGGCTGGGAGTACGCTGCAAGTTAAACTCAAGAAATTTGACGGG 895
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
DB 869 CGTGAAGTTGACCGGCTGGGAGTACGCTGCAAGATTAATAACTCAAGAAATTTGACGGG 928
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 896 GACCCCGCAACAGCGGTGAGTATGATGATTAATTCATGTAAGCCGAAAACTTACTTA 955
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 929 GACCCCGCAACAGCGGTGAGTATGATGATTAATTCATGTAAGCCGAAAACTTACTTA 988
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 956 CCCTTGACATGTAAGCGAATTTTCTAGAGATAGATAGTG--CTTCGGGAGCTTAACAG 1014
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 989 GTTTTGAATGATGCGGAGATCTTCGAGACGAGAGAGTGCCTTCGGAGCCGTAACAG 1048
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1015 GTGCTGATGAGCTGCTGCTCAGCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1074
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1049 GTGCTGATGAGCTGCTGCTCAGCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1108
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1075 GCAACCTTGTCAATTAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1134
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1109 GCAACCTTGTCAATTAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1168
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1135 AAACCGAGAGAGTGGGAGTGAACGTAAGTCCATGAGCCCTTAATGGGTAGAGGCTTAC 1194
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1169 AAACCGAGAGAGTGGGAGTGAACGTAAGTCCATGAGCCCTTAATGACGAGGCTTAC 1228
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1195 ACCTAATACATGAGCGGCTACAGAGGCTTCCAAACCGGAGGAGGAGGAGGAGGAGGAGG 1254
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1229 ACCTAATACATGAGCGGCTACAGAGGCTTCCAAACCGGAGGAGGAGGAGGAGGAGGAGG 1288
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1255 AGCGCGTGTAGTCCGATCGGATCGGAGTCTGCAACTGACTCCGTAAGTCCGAAATGCTAAGT 1314
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1289 AACGATCGTATGCTCGGATGCTGCAACTGCAATCGATGATGATGATGATGATGATGATGAT 1348
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1315 AATGCGGATACAGATGCTCCGCGTAAATAGTTCGCGGCTTGTACACACCGCGCTCA 1374
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1349 AATGCGGATACAGATGCTCCGCGTAAATAGTTCGCGGCTTGTACACACCGCGCTCA 1408
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1375 CACCATGAGAGTGGGTTTACACAGAACAGAGTGTAAACGTTAAGAGAGGCGCTTGCA 1434
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1409 CACCATGAGAGTGGGTTTACACAGAACAGAGTGTAAACGTTAAGAGAGGCGCTTGCA 1468
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1435 CGGTGAATTCATGACTGGGGTG 1457
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1469 CGGTGAATTCATGACTGGGGTG 1491
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 3  
US-08-114-695A-7  
; Sequence 7, Application US/08114695A  
; Patent No. 5508193  
; GENERAL INFORMATION:  
; APPLICANT: Mandelbaum, Raphael T.  
; APPLICANT: Wackelt, Lawrence P.  
; TITLE OF INVENTION: DEGRADATION OF S-TRIAZINES IN SOIL AND  
; TITLE OF INVENTION: WATER  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCHWEGMAN, LUNDBERG & WOESSNER, P.A.  
; STREET: 3500 IDS CENTER  
; CITY: MINNEAPOLIS  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/114,695A  
; FILING DATE: 31-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MUEITING, ANN M.  
; REGISTRATION NUMBER: 33,977  
; REFERENCE/DOCKET NUMBER: 600,268US1



QY 11 GCGCGATGCTTTACATCATGCAAGTCGACGCGATGCTTGCATCTGCTGCGA 70  
DB 11 GCGCGATGCTTTACATCATGCAAGTCGACGCGATGCTTGCATCTGCTGCGA 70  
QY 71 GTGGCGGATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 130  
DB 71 GTGGCGGATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 130  
QY 131 AGATGCTAATACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 190  
DB 131 AGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 190  
QY 191 TTTTGGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250  
DB 191 TTTTGGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250  
QY 251 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 310  
DB 251 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 310  
QY 311 CTACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 370  
DB 311 CTACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 370  
QY 371 GCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 430  
DB 371 GCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 430  
QY 431 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 490  
DB 431 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 490  
QY 491 CAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 550  
DB 491 CAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 550  
QY 551 CAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 610  
DB 551 CAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 610  
QY 611 ACTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670  
DB 611 ACTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670  
QY 671 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730  
DB 671 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730  
QY 731 CAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790  
DB 731 CAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790  
QY 791 TCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 850  
DB 791 TCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 850  
QY 851 CTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910  
DB 851 CTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910  
QY 911 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 970  
DB 911 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 970  
QY 971 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027  
DB 971 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027  
QY 1028 GTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087  
DB 1028 GTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087  
QY 1030 GTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089  
DB 1030 GTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089

QY 1088 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147  
DB 1090 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1144  
QY 1148 GTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207  
DB 1145 GTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1204  
QY 1208 GCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1267  
DB 1205 GCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264  
QY 1268 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1327  
DB 1265 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1324  
QY 1328 CATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1387  
DB 1325 CATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1384  
QY 1388 GGTTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1443  
DB 1385 GGTTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440

RESULT 5  
US-09-617-854A-1  
Sequence 1, Application US/09617854A  
Patent No. 6461840  
GENERAL INFORMATION:  
APPLICANT: BRAMUCCI, MICHAEL G.  
APPLICANT: MCCUTCHEN, CAROL M.  
APPLICANT: NAGARAJAN, VASANTHA  
APPLICANT: THOMAS, STUART M.  
TITLE OF INVENTION: Microbial Production of Terephthalic Acid and  
TITLE OF INVENTION: Isophthalic Acid  
FILE REFERENCE: CI-1257  
CURRENT APPLICATION NUMBER: US/09/617,854A  
CURRENT FILING DATE: 2001-09-04  
PRIOR APPLICATION NUMBER: 60/091,645  
PRIOR FILING DATE: JULY 2, 1998  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 1  
LENGTH: 1451  
TYPE: DNA  
ORGANISM: Burkholderia sp.  
US-09-617-854A-1

Query Match 75.8%; Score 1104.8; DB 3; Length 1451;  
Best Local Similarity 87.0%; Pred. No. 0;  
Matches 1250; Conservative 0; Mismatches 177; Indels 9; Gaps 3;

QY 11 GCGCGATGCTTTACATCATGCAAGTCGACGCGATGCTTGCATCTGCTGCGA 70  
DB 11 GCGCGATGCTTTACATCATGCAAGTCGACGCGATGCTTGCATCTGCTGCGA 70  
QY 71 GTGGCGGATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 130  
DB 71 GTGGCGGATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 130  
QY 131 AGATGCTAATACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 190  
DB 131 AGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 190  
QY 191 TTTTGGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250  
DB 191 TTTTGGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250  
QY 251 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 310  
DB 251 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 310

QY 311 CTACGGGAGGACAGTGGGGAATTTTGAACAATGGGCGACGCTGATCCAGCAATGCC 370  
DB 311 CTACGGGAGGACAGTGGGGAATTTTGAACAATGGGCGACGCTGATCCAGCAATGCC 370  
QY 371 GCGTAGTGAAGAGGCGCTTCGGGTTGTAAGCTCTTTCAGTGAAGAGGAGGTTACG 430  
DB 371 GCGTAGTGAAGAGGCGCTTCGGGTTGTAAGCTCTTTCAGTGAAGAGGAGGTTACG 430  
QY 431 GTAAATATATCGTGAACCATGACGGTATCGACAGAGAGCAACGGGCTTACTAGTCCAG 490  
DB 431 GTAAATATATCGTGAACCATGACGGTATCGACAGAGAGCAACGGGCTTACTAGTCCAG 490  
QY 491 CACCGCGGTAAATACGTAGGGGTGCAACGTTAACTCGAATTTCTGGGCGTAAAGGGGCG 550  
DB 491 CACCGCGGTAAATACGTAGGGGTGCAACGTTAACTCGAATTTCTGGGCGTAAAGGGGCG 550  
QY 551 CAGCGCGCTTGTAACTGAGTGAATCCCGGCGCTTAACTGGGAGTTGGCTTGA 610  
DB 551 CAGCGCGCTTGTAACTGAGTGAATCCCGGCGCTTAACTGGGAGTTGGCTTGA 610  
QY 611 ACTACAAAGCTTGAAGTGTGCAAGAGGAGGTGGAATTCATGTGTAGCACTGAATGCGT 670  
DB 611 ACTACAAAGCTTGAAGTGTGCAAGAGGAGGTGGAATTCATGTGTAGCACTGAATGCGT 670  
QY 671 AGAGATATGAGAGAACATCGATGGCGAAGGAGCGCTCGGGTTAACTAGAGCGTCATG 730  
DB 671 AGAGATATGAGAGAACATCGATGGCGAAGGAGCGCTCGGGTTAACTAGAGCGTCATG 730  
QY 731 CACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGTCACGCGCTTAAAGCATG 790  
DB 731 CACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGTCACGCGCTTAAAGCATG 790  
QY 791 TCACTAGATTGTTGGGCTTATTAAGCTTGTGTAACAGAGTCAACGCTGAATGTCACCG 850  
DB 791 TCACTAGATTGTTGGGCTTATTAAGCTTGTGTAACAGAGTCAACGCTGAATGTCACCG 850  
QY 851 CTGGGAGTACGCGTGGCAATTAATACTAAAGGAATTAACGAGGAGCGCGCAACGAGCG 910  
DB 851 CTGGGAGTACGCGTGGCAATTAATACTAAAGGAATTAACGAGGAGCGCGCAACGAGCG 910  
QY 911 TGGATATGATGATTAATTCGATGCAACGCGGAAACCTTACCTGACATGATGAC 970  
DB 911 TGGATATGATGATTAATTCGATGCAACGCGGAAACCTTACCTGACATGATGAC 970  
QY 971 GAATTTTCTGAAGATAGATTAGTCT---TCGGGAACGCTTAACACAGTGTCTGATGCT 1027  
DB 971 GAATTTTCTGAAGATAGATTAGTCT---TCGGGAACGCTTAACACAGTGTCTGATGCT 1027  
QY 1028 GTGTCAGCTGATGTGATGATGTTAGTCCCGCAACGAGGCGCAACCTTGTCA 1087  
DB 1028 GTGTCAGCTGATGTGATGATGTTAGTCCCGCAACGAGGCGCAACCTTGTCA 1087  
QY 1088 TTAATTCATCAATTTGGTGGCACTTTAATGAGATGTCGCGTGAACAAACGAGGAGAG 1147  
DB 1088 TTAATTCATCAATTTGGTGGCACTTTAATGAGATGTCGCGTGAACAAACGAGGAGAG 1147  
QY 1090 TTAGTTGC-----TAGCAAGAGCACTTAAGAGAGATCTCGCGGTGAACAAACGAGGAGAG 1144  
DB 1090 TTAGTTGC-----TAGCAAGAGCACTTAAGAGAGATCTCGCGGTGAACAAACGAGGAGAG 1144  
QY 1148 GTGGGAGTGAACGTCAGTCTCTCATGCGCTTATGGGATGAGGCTTCAACGATATACAAATG 1207  
DB 1148 GTGGGAGTGAACGTCAGTCTCTCATGCGCTTATGGGATGAGGCTTCAACGATATACAAATG 1207  
QY 1145 GTGGGAGTGAACGTCAGTCTCTCATGCGCTTATGGGATGAGGCTTCAACGATATACAAATG 1204  
DB 1145 GTGGGAGTGAACGTCAGTCTCTCATGCGCTTATGGGATGAGGCTTCAACGATATACAAATG 1204  
QY 1208 GCGCGTACAGAGGAGTGGCAACCGCGAGGAGGAGTGAATCTCAGAAAGGCGCGTGAAT 1267  
DB 1208 GCGCGTACAGAGGAGTGGCAACCGCGAGGAGGAGTGAATCTCAGAAAGGCGCGTGAAT 1267  
QY 1268 CCGGATCGAGTGTGCAACTCGACTCCGTGAAGTCCGGAATCGTGAATCGCGGATCAG 1327  
DB 1268 CCGGATCGAGTGTGCAACTCGACTCCGTGAAGTCCGGAATCGTGAATCGCGGATCAG 1327  
QY 1328 CATGTGCGGAGTGAATGCTTCCGGGCTTGTGACACACCGCGCTGACACATGAGGAGT 1387  
DB 1328 CATGTGCGGAGTGAATGCTTCCGGGCTTGTGACACACCGCGCTGACACATGAGGAGT 1387  
QY 1388 GGTTCACAGAGAGTGTAGTCTAATCGTAAGAGGAGGCGCTTGCACGCGTGAAT 1443  
DB 1388 GGTTCACAGAGAGTGTAGTCTAATCGTAAGAGGAGGCGCTTGCACGCGTGAAT 1443

Db 1385 GGTTCACAGAGAGTGTAGTCTAATCGTAAGAGGAGGCGCTTGCACGCGTGAAT 1440  
RESULT 6  
US-08-642-229A-3  
; Sequence 3, Application US/08642229A  
; Patent No. 5874291  
; GENERAL INFORMATION:  
; APPLICANT: Herwig, Russell P.  
; APPLICANT: Bielefeldt, Angela R.  
; APPLICANT: Stensel, H. David  
; APPLICANT: Strand, Stuart E.  
; TITLE OF INVENTION: Degradation of Environmental Toxins by a  
; TITLE OF INVENTION: Filamentous Bacterium  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Christensen O'Connor Johnson and Kindness PLLC  
; STREET: 1420 Fifth Avenue, Suite 2800  
; City: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: WA 98101-2347  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/642,229A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/246,865  
; FILING DATE: 20-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sheiness, Diana K.  
; REGISTRATION NUMBER: 35,356  
; REFERENCE/DOCKET INFORMATION:  
; TELEPHONE: (206) 682 8100  
; TELEFAX: (206) 224 0779  
; TELEX: 4938023  
; INFORMATION FOR SEQ. ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1455 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: nucleic acid  
; DESCRIPTION: "16S ribosomal DNA"  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Unknown. Possibly new species  
; STRAIN: A-1  
; US-08-642-229A-3  
Query Match 75.6%; Score 1101.8; DB 2; Length 1455;  
Best Local Similarity 86.9%; Pred. No. 0;  
Matches 1270; Conservative 1; Mismatches 179; Indels 12; Gaps 5;  
QY 1 ATTTGAACGCTGGCGGATGCTTTTACATGCAAGTTCGAAACGGCAGACGAGTCTTGCAAT 60  
DB 1 ATTTGAACGCTGGCGGATGCTTTTACATGCAAGTTCGAAACGGCAGACGAGTCTTGCAAT 60  
QY 61 CTGTGTCAGAGTGTGCGGAGCGGAGTGAATGATCGGAACGTAATCCAGAAAGAGGAGGAT 120  
DB 61 CTGTGTCAGAGTGTGCGGAGCGGAGTGAATGATCGGAACGTAATCCAGAAAGAGGAGGAT 120  
QY 59 CTGATGCGCAGTGTGCGGAGCGGAGTGAATGATCGGAACGTAATCCAGAAAGAGGAGGAT 118  
DB 59 CTGATGCGCAGTGTGCGGAGCGGAGTGAATGATCGGAACGTAATCCAGAAAGAGGAGGAT 118  
QY 121 ACGCATCGAAGATGTCTAATACCGCATATCTTAAGAGAGAAAGCAGGAGGATGAAA 180  
DB 121 ACGCATCGAAGATGTCTAATACCGCATATCTTAAGAGAGAAAGCAGGAGGATGAAA 180  
QY 119 ACTACTCGAAGAGTGTGAATACCGCATGATCTTAAGAGAGAAAGCAGGAGGATGCAA 178  
DB 119 ACTACTCGAAGAGTGTGAATACCGCATGATCTTAAGAGAGAAAGCAGGAGGATGCAA 178

181 GACCTTGCGCTTTTGGACGCGCGATGTCGATTAGCTAGTGGTGGGTTAAAGCGCTAC 240  
179 GACCTGCGGCTTACAGAGCGCGCTGGTGAGATTAGAGTGGTGGGATTAAGGCTTAC 238  
241 CAAGGCGACGATCATGTTGTTGGTTGAGAGAGCAGACGAG-CCACACTGGGATCTGAGAC 299  
239 CAAGCCACGATCTGTGCTGGTCTGAGAGAGCAGACGAGCCACACTGGGATCTGAGAGC 298  
300 GGCCGACGCTCTTACGAGGAGGAGCAGCTGGGGAATTTTGGACAATGGGCGGACCTGAT 359  
299 GGCCGACGCTCTTACGAGGAGGAGCAGCTGGGGAATTTTGGACAATGGGCGGACCTGAT 358  
360 CCAGCAATGCGCGCT-GAGTGAGAGAGGAGCTTGGGTTGTAAGCTTTTCACTGAGAA 418  
359 CCAGCAATGCGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 418  
419 GAAAGGTTACGGAATTAATCTGACCCATGACGCTATGACAGAGAGAGAGAGAGAGGAG 478  
419 GAAAGGCTCTCTTAATACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 478  
479 ACTACGTCAGAGCGCGCGGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 538  
479 ACTACGTCAGAGCGCGGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 538  
539 GTAAAGGAGTGGCAGGCGGCTTGTAAATGATGATGATGATGATGATGATGATGATGAT 598  
539 GTAAAGGAGTGGCAGGCGGCTTGTAAATGATGATGATGATGATGATGATGATGATGAT 598  
599 ATTGCGTTGAACTACAAAGCTGAGTGGCAGAGGAGGAGTGGATTCATGTTGAGC 658  
599 ACAGCGCTTGTGATGAGCAGAGGAGTGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 658  
659 AGTAAATGCGTGAATGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 718  
659 AGTAAATGCGTGAATGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 718  
719 CTGACGCTCATGACGAAAGCGTGGGAGAGCAAGATTAATGATGATGATGATGATGATGAT 778  
719 CTGACGCTCATGACGAAAGCGTGGGAGAGCAAGATTAATGATGATGATGATGATGATGAT 778  
779 CCCTAAAGCATGTCATGATGTTGGGCTTATTAAGCTTGTGATGATGATGATGATGATGAT 838  
779 CCCTAAAGCATGTCATGATGTTGGGCTTATTAAGCTTGTGATGATGATGATGATGATGAT 838  
839 GAAGTTGACCGCTGGGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 898  
839 GAAGTTGACCGCTGGGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 898  
899 CCGGACAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958  
899 CCGGACAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958  
959 TTGACATGAGCGAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1015  
959 TTGACATGAGCGAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1015  
999 TTGACATGAGCGAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1018  
1016 TGTGATGAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1075  
1019 TGTGATGAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078  
1076 CAACCTTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135  
1079 CAACCTTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133  
1136 AACCGAGAGAGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGAGGAGGAGGAGGAGGAG 1195  
1134 AACCGAGAGAGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGAGGAGGAGGAGGAGGAG 1193  
1136 CGTAAATCAATGCGGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1255  
1194 CGTAAATCAATGCGGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1253

1256 GCGCGTGTAGTCCGAGTGGAGTGTGACACTGCACTCCGTGAAGTCGGAATGCTAGTA 1315  
1254 GCCAGTGTAGTCCGAGTGGAGTGTGACACTGCACTCCGTGAAGTCGGAATGCTAGTA 1313  
1316 ATGCGGATCAGCATGTCGCGGTGAATACGTTCCGAGTCTTGTACACACCGCCGCTAC 1375  
1314 ATGCGGATCAGCATGTCGCGGTGAATACGTTCCGAGTCTTGTACACACCGCCGCTAC 1373  
1376 ACCATGGAGTGGGTTTACACAGAGAGGAGTGTACCTTAACGTAAGAGGAGGAGGAGGAG 1435  
1374 ACCATGGAGTGGGTTTACACAGAGAGGAGTGTACCTTAACGTAAGAGGAGGAGGAGGAG 1433  
1436 GGTGAATTCAGTACGCGGTG 1457  
1434 GCGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455

RESULT 7  
US-08-642-229A-2  
Sequence 2, Application US/08642229A  
Patent No. 5874291  
GENERAL INFORMATION:  
APPLICANT: Herwig, Russell P.  
APPLICANT: Bielefeldt, Angela R.  
APPLICANT: Stensel, H. David  
APPLICANT: Strand, Stuart E.  
TITLE OF INVENTION: Degradation of Environmental Toxins by a  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Christensen O'Connor Johnson and Kindness PLLC  
STREET: 1420 Fifth Avenue, Suite 2800  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: WA 98101-2347  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/642,229A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/246,865  
FILING DATE: 20-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sheiness, Diana K.  
REGISTRATION NUMBER: 35,356  
REFERENCE/DOCKET NUMBER: UOFW19233  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 682 8100  
TELEFAX: (206) 224 0779  
TELEX: 4938023  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1452 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
DESCRIPTION: "16S ribosomal DNA"  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Brachymonas denitrificans AS-P1  
US-08-642-229A-2

Query Match 75.6%; Score 1100.8; DB 2; Length 1452;  
Best Local Similarity 86.1%; Pred. No. 0;  
Matches 1257; Conservative 0; Mismatches 192; Indels 11; Gaps 3;

```
QY 1 ATTGAACGCTGGGCGCATGCTTTACACATGCAAGTCGAACGCGACGACGAGTGTTCAT 60
DB 1 ATTGAAGCTGGCGCGCATGCTTTACACATGCAAGTCGAACGCGTAAACGCT--CTTCGG 57
QY 61 CTGGTGGCGAGTGGCGGACGCGGTGATGATGCAATCGGAACGTATCCAGAAAGGCGGGTA 120
DB 58 ATGCTGACGAGTGGCGGACGCGGTGATGATGCAACGTGCGGACGCTGAGTGGGGATTA 117
QY 121 ACCGATCGAAAGATGTGCTAATACCGCATATCTTAAGAGGAAACGCGGGATCGAA 180
DB 118 ACTACTCGAAAGATGTGCTAATACCGCATGAACTGAGGTGGAACCGGGGCGCTTTG 177
QY 181 GACCTTGGCTTTTGGAGCGCGCGATGCTGATTAAGTACTAGTGGTGGGTAAGGCTTAC 240
DB 178 GGCCTCGCTACTGAGGCGCGCATATCAGATTAGGTGATGTTGGTGGGTAAGGCTTAC 237
QY 241 CAAGCGCAGCATGATGTTGCTGAGAGGACGACGACCACTGGGACTGAGACAG 300
DB 238 CAAGCGCAGCATGTTGATGCTGAGAGGACGACGACCACTGGGACTGAGACAG 297
QY 301 GCCCAGACTCTCTACGCGGAGGACGAGTGGGGAATTTTGGACATGGCGCAAGCTGATC 360
DB 298 GCCCAGACTCTCTACGCGGAGGACGAGTGGGGAATTTTGGACATGGCAAGCTGATC 357
QY 361 CAGCAATGCGCGCTGATGAAAGAGGCTTCGGGTTGTAAGCTCTTCAGTGGAGAGA 420
DB 358 CAGCAATGCGCGCTGAGAGGAGGAGGCTTCGGGTTGTAAGCTCTTCAGTGGAGAGA 417
QY 421 AAAGGTTACGGTAAATATGCTGACCCATGACGCTATCGACAGAGAGAGACCGGCTTAC 480
DB 418 AAAGGCTCTGTTAAATCTGGGCTCATGACGCTATCTGTAAGATTAAGACCGGCTTAC 477
QY 481 TAGCTGCGACGACCGCGGTATATCGTAGGTCGACGCTTATCGGAATTTACTGGGCT 540
DB 478 TAGCTGCGACGACCGCGGTATATCGTAGGTCGACGCTTATCGGAATTTACTGGGCT 537
QY 541 AAAGGTCGCGACGCGGCTTGTGTAAGTCAGATGGAATCCCGGGTTAAACCTGGGAT 600
DB 538 AAAGGTCGCGACGCGGCTTGTGTAAGTCAGATGGAATCCCGGGCTCAACCTGGGAT 597
QY 601 TGGCTTTGAAACTACAAAGCTAGAGTGGCAGAGGAGGTGGAATTCATGTTGACAG 660
DB 598 TGCATTGCTGATCTGCAAGCTGAGTGGCGGACGAGGGGATGGAATTCGCGGTGACAG 657
QY 661 TGAATGCTGATGATGAGATGAGAAACATCGATGCGAAGCGACCTCTGGGTTACACT 720
DB 658 TGAATGCTGATGATGAGATGAGAAACATCGATGCGAAGCGACCTCTGGGCTGCACT 717
QY 721 GAGGCTCATGACGAAAGCGTGGGAGACAAACGAGTTAATCCCTGGTATGTCACAGC 780
DB 718 GAGGCTCATGACGAAAGCGTGGGAGACAAACGAGTTAATCCCTGGTATGTCACAGC 777
QY 781 CTAACGATGCTCACTAGTGTGGGCTTATTAGCTTGGTGAAGAGCTTAAACGCTGA 840
DB 778 CTAACGATGCTCACTAGTGTGGGCTTATTAGCTTGGTGAAGAGCTTAAACGCTGA 837
QY 841 AGTTGACCGCTGGGAGGATGCGTGCAGATTTAAACTCAAGGAATTGACGGGACCC 900
DB 838 AGTTGACCGCTGGGAGGATGCGTGCAGATTTAACTCAAGGAATTGACGGGACCC 897
QY 901 GCAACAGCGGTGATTTATGAGATTAATTTGAGTCAACGCGAATAACCTTACCTTAC 960
DB 898 GCAACAGCGGTGATTTATGAGATTAATTTGAGTCAACGCGAATAACCTTACCTTAC 957
QY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC--TTCCGGAACGCTTACACAGTG 1017
DB 958 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC--TTCCGGAACGCTTACACAGTG 1017
QY 1018 CTGCATGGCTGTGCTCACTGCTGTGCTGAGATTTGGGTTAAGTCCCGACGAGCGCA 1077
DB 1018 CTGCATGGCTGTGCTCACTGCTGTGCTGAGATTTGGGTTAAGTCCCGACGAGCGCA 1077
```

```
QY 1078 ACCCTGTGATTAATATGATCATATTTGTTGGGCACTTTAATGAGACTGCCGTGACAA 1137
DB 1078 ACCCTGTGATTAATATGATCATATTTGTTGGGCACTTTAATGAGACTGCCGTGACAA 1132
QY 1138 CCGAGAAAGTGGGAGTGAACGCTCAAGCTTCAATGAGCTTTATGGTATGAGGCTTACAG 1197
DB 1133 CCGAGAAAGTGGGAGTGAACGCTCAAGCTTCAATGAGCTTTATGGTATGAGGCTTACAG 1192
QY 1198 TAATCAATGAGCGCTTACAGAGGTTTCCAAACCGCGAGGGGAGCTAATCTCAGAAAC 1257
DB 1193 TAATCAATGAGCGCTTACAGAGGTTTCCAAACCGCGAGGGGAGCTAATCTCAGAAAC 1252
QY 1258 GCGTGTAGTCCGATCGGATCGGATCTGCACTGCTCGTGAAGTCCGAATCGCTAGTAAT 1317
DB 1253 GCGTGTAGTCCGATCGGATCGGATCTGCACTGCTCGTGAAGTCCGAATCGCTAGTAAT 1312
QY 1318 CCGGATCAGCATGTGCGCGGTGAATCGTTCCCGGCTTTGTAACACCGCCGCTCACAC 1377
DB 1313 CCGGATCAGCATGTGCGCGGTGAATCGTTCCCGGCTTTGTAACACCGCCGCTCACAC 1372
QY 1378 CATGGAGTGGGTTTACAGAGAGGATGATCTAACCGTAAAGAGGGGCTTGCCACGG 1437
DB 1373 CATGGAGTGGGTTTACAGAGAGGATGATCTAACCGTAAAGAGGGGATACACCG 1432
QY 1438 TGAGATTGATGACTGGGGTG 1457
DB 1433 CAGGTTGCTGACTGGGGTG 1452

RESULT 8
US-09-735-567-4/c
; Sequence 4, Application US/09735567
; Patent No. 608190
; GENERAL INFORMATION:
; APPLICANT: Brameco, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BC1033 US NA
; CURRENT APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 1453
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Dinitrifying bacteria
; OTHER INFORMATION: Nucleotide sequence of rdna corresponding to strain RA2
US-09-735-567-4

Query Match 75.5%; Score 1100; DB 3; Length 1453;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 1256; Conservative 0; Mismatches 190; Indels 12; Gaps 3;

QY 1 ATTGAACGCTGGGCGCATGCTTTACACATGCAAGTCGAACGCGACGACGAGTGTTCAT 60
DB 1449 ATTGAAGCTGGCGCGCATGCTTTACACATGCAAGTCGAACGCGACGCGGCGCA----AC 1394
QY 61 CTGGTGGCGAGTGGCGGACGCGGTGATGATGCAATCGGAACGTATCCAGAAAGGCGGGTA 120
DB 1393 CTGGTGGCGAGTGGCGGACGCGGTGATGATGCAATCGGAACGTATCCAGAAAGTGGGGATTA 1334
QY 121 ACCGATCGAAAGATGTGCTAATACCGCATATCTTAAGAGGAAACGCGGATCGAA 180
DB 1333 GCCCGCGAAAGCGGATTAATACCGCATGTGATGAGATGAAAGTGGGCGACGCGCA 1274
QY 181 GACCTTGGCTTTTGGAGCGCGCGATGCTGATTAAGTGGTGGGTTAAAGGCTTAC 240
DB 1273 GGCCTCGCTACTGAGGCGCGCATATCAGATTAGGTGATGTTGGTGGGTTAAAGGCTTAC 1214
```

```
OY 241 CAAGCGACATCATAGTGTGCTGAGAGAGACCACTGGACCTGAGACAG 300
    |||
DB 1213 CAAGCTGGAGATGTAGTGTGCTGAGAGATGATCAGCACTGGAGCTGAGACAG 1154
OY 301 GCCCAGACTCTTACGGGAGGACAGCTGGGAAATTTTGAACAATGGGCCCAAGCTGATC 360
    |||
DB 1153 GCCCAGACTCTTACGGGAGGACAGCTGGGAAATTTTGAACAATGGGCCCAAGCTGATC 1094
OY 361 CAGCATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGAAGA 420
    |||
DB 1093 CAGCCATTCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGAAGA 1034
OY 421 AAGGTTACGGTAATATCTGTGACCATGACGGTATCGACAGAGAAAGCACCGGCTAAC 480
    |||
DB 1033 AAGGCTCTCTTAATACAGGGGGCATATGACGGTACCGTAAGATTAAGCACCGGCTAAC 974
OY 481 TACGTGCACAGACCCCGGTAAATCTGTAAGGTGCAACCGTTAATCGGAATTAACCTGGCCGT 540
    |||
DB 973 TACGTGCACAGACCCCGGTAAATCTGTAAGGTGCAACCGTTAATCGGAATTAACCTGGCCGT 914
OY 541 AAGGGGTGCGACGGCGCTTGTAAATGATGATGTAATCCCGGCTTAACCTGGGAAT 600
    |||
DB 913 AAGCGGTGCGACGGCGCTTGTAAATGATGATGTAATCCCGGCTTAACCTGGGAAT 854
OY 601 TGGCTTTGAAACTACAAAGCTAGAGTGTGACAGAGGAGGTGAATTCATGTGTAGCAG 660
    |||
DB 853 TGGCTTTGTAAGTGTGAAAGCTGAGGTGCGGACAGAGGAGGTGAATTCATGTGTAGCAG 794
OY 661 TGAATTCGTAGATATGGAAGAACTCGATGCGGACAGGCGCTCTGGGTTAACACT 720
    |||
DB 793 TGAATTCGTAGATATGGAAGAACTCGATGCGGACAGGCGCTCTGGGTTAACACT 734
OY 721 GAGGCTATACAGAAAGCGTGGGAGACAAACAGGATTAAGTACCTGTAGTCAAGCC 780
    |||
DB 733 GAGGCTATACAGAAAGCGTGGGAGACAAACAGGATTAAGTACCTGTAGTCAAGCC 674
OY 781 CTAAACGATGCACTAGTGTGGGCTTAATTAAGCTGTGTAACGAGCTAACGCGTGA 840
    |||
DB 673 CTAAACGATGCACTAGTGTGGGCTTAATTAAGCTGTGTAACGAGCTAACGCGTGA 614
OY 841 AGTTGACCGCTGGGAGATACGCTGCGAAGTTAAACTCAAAGAAATGACGGGAGCC 900
    |||
DB 613 AGTTGACCGCTGGGAGATACGCTGCGAAGTTAAACTCAAAGAAATGACGGGAGCC 554
OY 901 GCACAAAGCGGTGATTAATGCTGATTAATGCTGCAACGCGGAAACCTTACTACCTT 960
    |||
DB 553 GCACAAAGCGGTGATTAATGCTGATTAATGCTGCAACGCGGAAACCTTACTACCTT 494
OY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTCT--TCGGGAAACGCTTAACAAGGTG 1017
    |||
DB 493 GACATGTAGCGAATTTTCTAGAGATAGATTAGTCT--TCGGGAAACGCTTAACAAGGTG 434
OY 1018 CTGATGAGCTGTGCTGACGCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1077
    |||
DB 433 CTGATGAGCTGTGCTGACGCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 374
OY 1078 ACCCTTGCATTAATGCTGATCTATTTGGTGGCACTTAATGAGTCCCGGTGACAAA 1137
    |||
DB 373 ACCCTTGCATTAATGCTGATCTATTTGGTGGCACTTAATGAGTCCCGGTGACAAA 319
OY 1138 CCGAGAGAAAGTGGGAGTGAAGTCTCAATGCTCAATGAGGCTTAATGAGGCTTCAACAG 1197
    |||
DB 318 CCGAGAGAAAGTGGGAGTGAAGTCTCAATGCTCAATGAGGCTTAATGAGGCTTCAACAG 259
OY 1198 TAATCAATGAGCGCTGACAGAGGTTGCGAACCGCGAGAGGAGACTAATCTCAGAAAGC 1257
    |||
DB 258 TAATCAATGAGCGCTGACAGAGGTTGCGAACCGCGAGAGGAGACTAATCTCAGAAAGC 199
OY 1258 GCGTGTATGCTGGATTCGAGTCTGCAACTCGGCTGGAAGTGGGAATGCTGTATAT 1317
    |||
DB 198 GCGTGTATGCTGGATTCGAGTCTGCAACTCGGCTGGAAGTGGGAATGCTGTATAT 139
OY 1318 GCGGATACGATGTGCGGCTGAATACGTTCCGGGTCTTGTACACACCGCCCTGACAGC 1377
```

```
DB 138 CGTGATACGATGCTCAGCGTGAATAGTTCGCCGCTCTTGTAACACACCGCCCTCACAC 79
    |||
OY 1378 CATGGAGTGGGTTTACACAGAGAGAGTACTCTAACCGTAAGAGGGCGCTTGGCACAG 1437
    |||
DB 78 CATGGAGAGGGGTTCTGCCAGAAATGATTAGCTTAACCGCAAGAGGGCGATTAACACAG 19
    |||
OY 1438 TGAGATTCAATGACTGGG 1455
    |||
DB 18 CAGGGTTCGTGACTGAGG 1
    |||

RESULT 9
US-08-114-695A-8
; Sequence 8, Application US/08114695A
; Patent No. 5508193
; GENERAL INFORMATION:
; APPLICANT: Mandelbaum, Raphael T.
; APPLICANT: Mackett, Lawrence P.
; TITLE OF INVENTION: WATER
; TITLE OF INVENTION: WATER
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCHUEGMAN, LUNDBERG & MOESSNER, P.A.
; STREET: 3500 IDS CENTER
; CITY: MINNEAPOLIS
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Releasee #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/114,695A
; FILING DATE: 31-Aug-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MUEITING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 600,268US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: rRNA
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas cepacia
; US-08-114-695A-8

Query Match 74.9%; Score 1091; DB 2; Length 1474;
Best Local Similarity 68.3%; Pred. No. 0;
Matches 991; Conservative 247; Mismatches 205; Indels 9; Gaps 3;

OY 1 ATTAAGCGTGGCGGATGCTTATACATCAAGTGAAGGCG-AGCAGGATGCTTGA 59
    |||
DB 28 AUNAAAGCGUGGCGGACUGGCUUACACAUAGUCGAAAGCGACAGCAUGGGGUGCUU 87
    |||
OY 60 TCTGGTGGCGAGTGGCGGACGGGTGAGTAAATGATCGAAGCTATCCAGAAAGGGGGGT 119
    |||
DB 88 ACCUGGCGCGAGUGGCGAAGCGGUGAGUUAUACUGCGAACAUUGUCCUGAUGUGGGGAGU 147
    |||
OY 120 AACGCATCGAAAGATGCTAATACCGCATTAATCTTAAGAGAGAAAGCGAGGATCGAA 179
    |||
DB 148 AGCAGCGGAAAGCGCNAUUAUACCGCAUACGACAUUCAAAGGCGGAGACCUUC 207
    |||
OY 180 AGACCTTGCGCTTTTGAAGCGGCGGATGCTGATTAAGTGTGGGGTAAAGGCTTA 239
```



Db 208 GGGCCUCGCCUUAUAGGUGUGCCGAUUGGUGUAUUAUCUUAUGUGUGGUAUAGGCCUA 267  
Qy 240 CCAAGGCGAGATCAAGTGTGTCTGAGAGAGACAGACCACTGGGACTGAGACAC 239  
Db 268 CCAAGGCGAGATCAAGTGTGTCTGAGAGAGAGACAGACCACTGGGACTGAGACAC 327  
Qy 300 GGGCCGAGATCTCAACGGGAGAGAGAGATGAGGAAATTTGGACAATGGGCGCAAGCTGAT 359  
Db 328 GGGCCGAGATCTCAACGGGAGAGAGAGATGAGGAAATTTGGACAATGGGCGCAAGCTGAT 387  
Qy 360 CCAGCAATGCGCGCTGAGTGAAGAGAGCTTCCGGTGTGAAGCTCTTTCAGTCCGAGAG 419  
Db 388 CCAGCAATGCGCGCTGAGTGAAGAGAGCTTCCGGTGTGAAGCTCTTTCAGTCCGAGAG 447  
Qy 420 AAAAGGTTACGTTAAATTAATCTGACCCATGACCGTATCGACAGAAAGACCGGCTTA 479  
Db 448 AAUUCUUGGCUUAUUAACAGCCGGGGAGUAGCGUUAACGGAAGAAUUAAGCACCGGCUA 507  
Qy 480 CTAAGTGGCGAGAGCGCGGCTTAATAGTGGGCAAGCTTAATCGGAATTACTGGGCG 539  
Db 508 CUAAGGCGAGAGCGCGGCTTAATAGTGGGCAAGCTTAATCGGAATTACTGGGCG 567  
Qy 540 TAAAGGCTGCGAGAGCGCGCTTGTAACTGATGTAATCCCGGGCTTAACCTGGGAA 599  
Db 568 UAAAGGCTGCGAGAGCGCGCTTGTAACTGATGTAATCCCGGGCTTAACCTGGGAA 627  
Qy 600 TTGCGTTTGAACCTAACAGCTAGTGTGCGAGAGAGAGTGAATTCATCTGTAGCA 659  
Db 628 CUGCAUUGUGUAGCUGGACAGGCUAGAUUENNAAGGGGGUAGAAUUAACACUGUAGCA 687  
Qy 660 GTGAATAGCTTAAGATATGGAAGAACTAGATGGGAGAGAGCGCTCCGGGCTTAACAC 719  
Db 688 GUAAGAAUUGGCUUAUUAAGUAGAGAAUUAACGAGAGAGAGCGCTCCGGGCTTAACAC 747  
Qy 720 TGAAGCTCAAGCAAGAAAGCTGGGAGCAAAAGAGATTTAGATACCTGGTACTGACAGC 779  
Db 748 UGACGCUCAUGCAAGAAAGCGUGGAGAGCAAAAGAGUUAAGUUAUCCUGUUAUCCACAGC 807  
Qy 780 CCTAAACGATGTCAACTAGTTGTTGGGCTTAATTAAGGCTTGTGAAGCAAGCAAGCTGAC 839  
Db 808 CUAAGAAAGUUAUUAAGUAGAGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 867  
Qy 840 AAGTTGACCGCTGGGAGAGTACGCTCGCAAGATTTAAACCTCAAGAAATTTAGAGGAGAC 899  
Db 868 AAGTTGACCGCTGGGAGAGTACGCTCGCAAGATTTAAACCTCAAGAAATTTAGAGGAGAC 927  
Qy 900 CGCAGAGCGGTGATTTATGTGATTTAATCGATGCAACGCGAAAAAATTCTTACCTGCT 959  
Db 928 CGCAGAGCGGTGATTTATGTGATTTAATCGATGCAACGCGAAAAAATTCTTACCTGCT 987  
Qy 960 TGAATGTAGCGAAATTTCTAGAGATAGATTAGTGTCT--TCGGAGACGCTTAACACAGCT 1016  
Db 988 UGACUUGUGCGAAUUCUGUGAGAGCGGGAGUGUCUGAAAGAGAAACGCGCGCACAGGU 1047  
Qy 1017 GCTGATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076  
Db 1048 GCTGATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1107  
Qy 1077 AACCTTGTCAATTTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1136  
Db 1108 AACCTTGTCAATTTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1162  
Qy 1137 ACCGAGAGAGAGTGGAGATGACGTCAAGTCTCATGGGCTTATGGGTAGGCTTCAAC 1196  
Db 1163 ACCGAGAGAGAGTGGAGATGACGTCAAGTCTCATGGGCTTATGGGTAGGCTTCAAC 1222  
Qy 1197 GTATATCAATGCGCGCTGACAGAGGTTGCCAACCCGCGAGAGGAGAGTATCTCAAGAG 1256  
Db 1223 GTATATCAATGCGCGCTGACAGAGGTTGCCAACCCGCGAGAGGAGAGTATCTCAAGAG 1282  
Qy 1257 CGGCTGTAGTCCGAGTCCGAGTCTGCAACTCGGAGCTCCGGAAGTGGGAATCGCTAGTA 1316

Db 1283 CCGAUCGUAUUGCCGGAUUGCACUCUGCAACUCGAGUGAGUAGAGUGGAAUUCGUAUUA 1342  
Qy 1317 TGGCGGATTCAGATGTCGCGGAGAAATACGTTCCGGGTCTGTACACACCGCGGTGACA 1376  
Db 1343 UCGCGGAUAGCAUUGCCGCGGAGAAUAGCUUCCCGGAGUUGUUAACAGCAGCCGCUACA 1402  
Qy 1377 CCATGGAGTGGGTTCACAGAAACAGTACTTAAACGTTAAGAGAGGCGCTTCCAGC 1436  
Db 1403 CCAUGGAGUGGUGUUAUUAACAGAAUUGGUGUUAUUAUUAUUAUUAUUAUUAUUAUUA 1462  
Qy 1437 GTGAGATTCATG 1448  
Db 1463 GUAGAUUUAAG 1474

RESULT 10  
US-09-735-567-6/c  
; Sequence 6, Application US/09735567  
; Patent No. 6608190  
; GENERAL INFORMATION:  
; APPLICANT: Brannucci, Michael  
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
; FILE OF INVENTION: Industrial Wastewater Bioreactors  
; FILE REFERENCE: BC1033 US NA  
; CURRENT APPLICATION NUMBER: US/09/735,567  
; PRIOR FILING DATE: 2000-12-13  
; PRIOR APPLICATION NUMBER: 60/1171,140  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 1539  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Brachyomonas  
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RA6  
US-09-735-567-6

Query Match 74.2%; Score 1080.8; DB 3; Length 1539;  
Best Local Similarity 85.7%; Pred. No. 0;  
Matches 1251; Conservative 0; Mismatches 197; Indels 12; Gaps 4;

Qy 3 TGAAGCGTGGGCGGAGCTTTTACATGCAAGTGAAGCGGAGCGAGTGTGATCT 62  
Db 1538 TGAAGCGTGGGCGGAGCTTTTACATGCAAGTGAAGCGGAGCGAGTGTGATCT 1481  
Qy 63 GGTGGCGAGTGGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 122  
Db 1480 GGTGGCGAGTGGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421  
Qy 123 GCATGAAAGATGTGCTAATCCGATATCTTAAAGAGAAAGCAGAGGATC--GAAA 180  
Db 1420 TCGGGGAAAGCGGATTAATACCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361  
Qy 181 GACCTTGGCTTTTGAAGGCGCGATGTCTGATTTAGTGTGATGATGATGATGATGATGATGAT 240  
Db 1360 GACCTTGGCTTTTGAAGGCGCGATGTCTGATTTAGTGTGATGATGATGATGATGATGATGAT 1301  
Qy 241 CAAAGCGAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
Db 1300 CAAAGCGAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1241  
Qy 301 GCCCAGACTCTTACGCGAGAGAGCAGATGGGGAATTTTGAACAATGGGCGCAAGCTGATC 360  
Db 1240 GCCCAGACTCTTACGCGAGAGAGCAGATGGGGAATTTTGAACAATGGGCGCAAGCTGATC 1181  
Qy 361 CAGCATGCGCGGTGAGGAGAAAGCCCTTGGGTGATTAAGCTTTTCAATGCGAGAA 420  
Db 1180 CAGCATGCGCGGTGAGGAGAAAGCCCTTGGGTGATTAAGCTTTTCAATGCGAGAA 1121  
Qy 421 AAAGGTTACGGAATTAATCGTGAACCATGACGATGACAGAGAAAGCAACCGGCTAAC 480



Db	1120	AAGGGCTCTTTCTTAATAAAGGGCCACATGACGGTACCCTGAAGATTAAGCACGGGCTAAC	1061
Qy	481	TACGTGCGACAGCCCGCGTAATATGAGGGTGCAGCGTTAATCGGAATTACTGGGCGT	540
Db	1060	TACGTGCGACAGCCCGCGTAATACGTAGGGTGCAGCGTTAATCGGAATTACTGGGCGT	1001
Qy	541	AAAGGGTGCAGGGCGGCTTTGTAATGTCAGATGTGAATCCCGGGCTTAACTCGGGAAAT	600
Db	1000	AAAGCGTGCAGGGCGGCTTTGTAAGACAGATGTGAATCCCGGGCTTAACTCGGGAAAC	941
Qy	601	TGCGTTTGAACCTCAAAAGCTAGAGTGTGGCAGAGGGAGGTGAATTCAATGTGTACAG	660
Db	940	TGCCATTGTGACTCAAAAGCTGAGATACGCGAGAGGGGAGTGAATTCGCCGTGTACAG	881
Qy	661	TGAATCGCTAGAGATATGGAAGAACATCGATGCGAAGGACGCTCTCGGTTAACACT	720
Db	880	TGAATCGCTAGATATGCGAGGAACAACCGATGCGAAGGACATCCCTGGGCTGTAACT	821
Qy	721	GACGCTCATGCAAGAAAGCTGGGGAGCAACAGATTAGTATACCTGTGTGTCAAGCC	780
Db	820	GACGCTCATGCAAGAAAGCTGGGGAGCAACAGATTAGTATACCTGTGTGTCAAGCC	761
Qy	781	CTAAACGATGCATCTAGTGTGTGGGCGCTTATTAAGGCTGTGTAAGAACTTAACGCGTGA	840
Db	760	CTAAACGATGCATCTAGTGTGTGGGCGCTTATTAAGGCTGTGTAAGAACTTAACGCGTGA	701
Qy	841	AGTTGACCGCGTGGGGAGTACGGTTCGACAGATTAAACTCAAGGAATTGACGGGACCC	900
Db	700	AGTTGACCGCGTGGGGAGTACGGGCGCAAGGTTAAACTCAAGGAATTGACGGGACCC	641
Qy	901	GCACAAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAATAAACCCTTAACCTT	960
Db	640	GCACAAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAATAAACCCTTAACCTT	581
Qy	961	GACATGTAGGCAATTTTCTAGAGATAGATTATGTGCT---TCGGGAAAGCTTAACAGAGTG	1017
Db	580	GACATGTATGGAATCCCGAGAGATGTGGAGTGTCTCGCAAGAGAGCAATAACAGAGTG	521
Qy	1018	CTGCATGCTGTCTGCACTCGTGTCTGTGAAGATGTTGGGTTAATGTCGCGCAACGAGCGCA	1077
Db	520	CTGCATGCTGTCTGCACTCGTGTCTGTGAAGATGTTGGGTTAATGTCGCGCAACGAGCGCA	461
Qy	1078	AACCTTGTCATTAATTGCGCATCATTTGGTGTGGGCACTTAAATGAGACTGCCGCTGACAA	1137
Db	460	AACCTTGCGCATCATGTTGCTACGAAA-----GGGCACTGTGAATGGACTGCCGCTGACAAA	406
Qy	1138	CCGAGGAAAGGTGGGGATGAGCTCAAGTCTCATAGGACCCCTTAATGGGTATGGGCTTACACG	1197
Db	405	CCGAGGAAAGGTGGGGATGAGCTCAAGTCTCATAGGACCCCTTAATAGGTGGGCTTACACG	346
Qy	1198	TAAATCAATGCGCGCTTACAGAGGTTGCGCAACCCGCGAGGGGAGCTAATCTCAGAAAGC	1257
Db	345	TCATACAAATGGCGCGTACAAAGGGCAGCGAAGCCCGAGGTGAAACCAATCCCATTAAGC	286
Qy	1258	GGGTGTAGTCCGATTCGAGTCTGCAACTCGACTTCGTAAGTCCGTAATGCTATGTAAT	1317
Db	285	CGGTGTAGTCCGATTCGAGTCTGCAACTCGACTTCGTAAGTCCGTAATGCTATGTAAT	226
Qy	1318	CGCGGATCAGCAATGTCGCGGTAAATACGTTCCCGGGTCTTGTACACAACGCGCCGTCACAC	1377
Db	225	CGGTGATCAGCAATGTCACGCTAAATACGTTCCCGGGTCTTGTACACAACGCGCCGTCACAC	166
Qy	1378	CATGGGAGTGGGTTTCAACAGAGAGGATGCTTAACGTTAAGAGAGGCGCTTGCACGG	1437
Db	165	CATGGGAGGCGGTTCTACACAGAGCAGATGACCTTAACGCAAGAGAGGCGCTTGCACGG	106
Qy	1438	TGAGATTCAATGACTGGGCTG 1457	
Db	105	TGGGCTTCTGACTGGGCTG 86	

```

: Sequence 7, Application US/09735567
: Patent No. 6608190
: GENERAL INFORMATION:
: APPLICANT: Brammucci, Michael
: TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
: TITLE OF INVENTION: Industrial Wastewater Bioreactors
: FILE REFERENCE: BC1033 US NA
: CURRENT APPLICATION NUMBER: US/09/735,567
: CURRENT FILING DATE: 2000-12-13
: PRIOR APPLICATION NUMBER: 60/117,140
: PRIOR FILING DATE: 16 DECEMBER 1999
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 7
: LENGTH: 1454
: TYPE: DNA
: ORGANISM: Unknown Organism
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: Brachyomonas
: OTHER INFORMATION: Nucleotide sequence of the 16S rDNA corresponding to strain RA9
: US-09-735-567-7

```

Query Match	74.1%	Score 1080.2	DB 3	Length 1454
Best Local Similarity	85.6%	Pred. No. 0	Mismatches 198	Indels 12
Matches 1251	Conservative	0	Pident	Gaps 4
QY	2	TTGAACGCTGGCGGCATGCTTTACACATGCAATGCAAGTCGAAACGGCAGCAGATGCTTTCATC	61	
Db	1	TTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAAACGGCAGCAGATGCTTTCATC	58	
QY	62	TGTTGGCGATGGCGGACGGGTTGATATGTCATTCGAAACGTTATCCAGAAAGGGGGCTTA	121	
Db	59	TGGCGCGAGTGGCGAAACGGGTTGATTAAGCATCGAAACGGGCCCGGTATGGCGGATAG	118	
QY	122	CGCATCCGAAATATGCTTAATACCGCATATATCTGTAAGGAGAAAGCAGGAGA--TCGAA	179	
Db	119	CTGGCGGAAACCGGATTAATACCGCATAGATCCGTGATGAAACAGGGGATCTCGCA	178	
QY	180	AGACCTTGCGCTTTTGGAGCGCGCGATGCTGATTAAGCTAGTTGGTGGGTAAAGGCTTA	239	
Db	179	GGGCTTGCGCTTACGTGAGAGCGGCGCATGTGATTAAGTATGATGTTGGTGGGTAAAGGCCA	238	
QY	240	CCAAGGCGCATCATGATGTTGTTGTTGAGAGAGACCAAGCCACACTGGGACTAGAACAC	299	
Db	239	CCAAGCTCGCATCTGATGCTGTGCTGTGAGAGATGATACCAACTCGGACTGAGAAC	298	
QY	300	GGCCCAAGCTCTTACGGGAGGACAGAGTGGGAAATTTTGAACAATGGGCGCAAGCCTGAT	359	
Db	299	GGCCCAAGCTCTTACGGGAGGACAGAGTGGGAAATTTTGAACAATGGGCGCAAGCCTGAT	358	
QY	360	CCAGCAATGCGCGCTGATGTAAGAAAGGCTCTCGGGTTGTAAGCTCTTTCAGTCGAGAG	419	
Db	359	CCAGCAATGCGCGCTGATGTAAGAAAGGCTCTCGGGTTGTAAGCTCTTTCAGTCGAGAG	418	
QY	420	AAAAGTTACGGTAAATATATGTCGACCCATGACGGTATCGACAGAGAAGCAACGGGCTTA	479	
Db	419	AAAAGGCTCTTCTTAATAAGAGGACATGACGGTATCCGTAAAGATTAAGCAACGGGCTTA	478	
QY	480	CTACGTCCACAGACGCCGGGCTTAATCGTAAGGTCGCAACGGTTAATCCGAAATTACTGGGCG	539	
Db	479	CTACGTCCACAGACGCCGGGCTTAATCGTAAGGTCGCAACGGTTAATCCGAAATTACTGGGCG	538	
QY	540	TAAAGGATGCGCAGCGGCTTGTATGATCAGATGTGAATCCCGGGGCTTAACCTTGGAA	599	
Db	539	TAAAGGATGCGCAGCGGCTTGTGCAAGACATGTGTGAATCCCGGGGCTCAACCTTGGAA	598	
QY	600	TTGCGCTTTGAAACTTACAAAGCTAGAGTGTGCGACAGAGGAGTGTGAATTCTATGTGTACCA	659	
Db	599	TTGCGCTTTGTGACTCAAGAGCTGTGAGTACGCGAGAGGGGAGATGTGAATTCTCGGTGTACCA	658	
QY	660	GTGAAGATCGTGAATATGGAAGAAACATGATGGCGAAAGGACGCTCTGGGTTTAACAC	719	
Db	659	GTGAAGATGTGATATGCGAGAGAAACCATGTGCGAAAGGACATGCTCCCTTGGGCTTGTAC	718	

QY 720 TGAACGTCATGCAAGCAAGCGTGGGAGCAAAAGAGATTAGATACCGTGTAGTCCAGCC 779  
DB 719 TGAACGTCATGCAAGCAAGCGTGGGAGCAAAAGAGATTAGATACCGTGTAGTCCAGCC 778  
QY 780 CCTAAACGATGTCATGATGTTGGGACCTTATTAGAGCTTGTGTAACGAAGCTAACCGGTG 839  
DB 779 CCTAAACGATGTCATGATGTTGGGAGCTTATTAGAGCTTGTGTAACGAAGCTAACCGGTG 838  
QY 840 AAGTTGACCGGCTGGGAGATACGGTTCGCAAGATTAAACTCAAGGAATTGACGGGAGCC 899  
DB 839 AAGTTGACCGGCTGGGAGATACGGTTCGCAAGATTAAACTCAAGGAATTGACGGGAGCC 898  
QY 900 CGCACAAGCGGTGATGATGTTGATTAATTCGATGCAACGCGGAAAAAACCCTTACCTACCT 959  
DB 899 CGCACAAGCGGTGATGATGTTGATTAATTCGATGCAACGCGGAAAAAACCCTTACCTACCT 958  
QY 960 TGACATGATGCAATTTTCTAGAGATTAGATTAGTCTCTCTCGGAAACGCTAACACAGGT 1016  
DB 959 TGACATGATGCAATTTTCTAGAGATTAGATTAGTCTCTCGGAAACGCTAACACAGGT 1018  
QY 1017 GCTGATGAGCTGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076  
DB 1019 GCTGATGAGCTGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078  
QY 1077 AACCTTGTCTATTAATTTGCAATCTTGTGTTGGGCACTTAAATGAGACTGCGGTGACAA 1136  
DB 1079 AACCTTGTCTATTAATTTGCAATCTTGTGTTGGGCACTTAAATGAGACTGCGGTGACAA 1133  
QY 1137 ACCGAGGAAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196  
DB 1134 ACCGAGGAAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193  
QY 1197 GTAATCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256  
DB 1194 GTAATCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253  
QY 1257 CGGCTGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1316  
DB 1254 CGGCTGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313  
QY 1317 TCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1376  
DB 1314 TCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373  
QY 1377 CCATGGAAGTGGGTTTCAACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1436  
DB 1374 CCATGGAAGTGGGTTTCAACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1433  
QY 1437 GTGAGATTCAATGCTGGGGTG 1457  
DB 1434 GTGAGATTCAATGCTGGGGTG 1454

RESULT 12  
US-09-735-567-2/c  
Sequence 2, Application US/09735567  
Patent No. 6608190  
GENERAL INFORMATION:  
APPLICANT: Brumccj, Michael  
TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
FILE REFERENCE: BC1033 US NA  
CURRENT APPLICATION NUMBER: US/09/735,567  
PRIOR FILING DATE: 2000-12-13  
PRIOR APPLICATION NUMBER: 60/171,140  
PRIOR FILING DATE: 16 DECEMBER 1999  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 2  
LENGTH: 1467  
TYPE: DNA  
ORGANISM: Unknown Organism

FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Brachymonas  
OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain NBE13  
US-09-735-567-2  
Query Match 74.1%; Score 1079.6; DB 3; Length 1467;  
Best Local Similarity 85.6%; Pred. No. 0;  
Matches 1251; Conservative 0; Mismatches 199; Indels 12; Gaps 4;  
QY 1 ATTTGAACGCTGGCGGCTGCTTATACATGCAATGCAAGCGGACACGAGCTTGCTGAT 60  
DB 1465 ATTTGAACGCTGGCGGCTGCTTATACATGCAATGCAAGCGGACACGAGCTTGCTGAT 1408  
QY 61 CTGGTGGCGAGTGGCGGACGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 1407 ATGGCGGCGAGTGGCGGACGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1348  
QY 121 ACGCATGCAAGATGCTATATACCGCATATATCTTAAAGAGAAAAGCAGGGGA--TCGA 178  
DB 1347 GCTCGGCGAAAACCGGATTTATATACCGCATGATCCTGTGATGAAAGCAGGGGACTCGCA 1288  
QY 179 AAGACTTGGCTGCTTTTGGAGCGGCGGATGCTGATGATGATGATGATGATGATGATGATGATGAT 238  
DB 1287 AAGGCTTGGCTGCTTTTGGAGCGGCGGATGCTGATGATGATGATGATGATGATGATGATGATGAT 1228  
QY 239 ACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298  
DB 1227 ACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1168  
QY 299 CGGCGCGACGCTCTACGCGGAGCAGCAGTGGGAAATTTTGACAAATGGCGCAACGCTGTA 358  
DB 1167 CGGCGCGACGCTCTACGCGGAGCAGCAGTGGGAAATTTTGACAAATGGCGCAACGCTGTA 1108  
QY 359 TCCAGCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418  
DB 1107 TCCAGCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1048  
QY 419 GAAAAGGTAAAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 478  
DB 1047 GAAAAGGTAAAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 988  
QY 479 ACTAGTGGCAGCAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538  
DB 987 ACTAGTGGCAGCAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 928  
QY 539 GTTAAAGGTGCGGAGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 598  
DB 927 GTTAAAGGTGCGGAGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 868  
QY 599 ATTGCGTTGAAACTTACAAAGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 658  
DB 867 ATTGCGTTGAAACTTACAAAGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 808  
QY 659 AGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
DB 807 AGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748  
QY 719 CTGACGCTATGACGAAAGCGTGGGAGCAAAAGATTAATGATGATGATGATGATGATGATGATGAT 778  
DB 747 CTGACGCTATGACGAAAGCGTGGGAGCAAAAGATTAATGATGATGATGATGATGATGATGATGAT 688  
QY 779 CCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838  
DB 687 CCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628  
QY 839 GAAAGTTGACCGGCTGGGAGTACCGTTCGCAAGATTAAATCTCAAGGAATTGAACGGGAGC 898  
DB 627 GAAAGTTGACCGGCTGGGAGTACCGTTCGCAAGATTAAATCTCAAGGAATTGAACGGGAGC 568  
QY 899 CGGACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958  
DB 567 CGGACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508

QY	955	TTGACATGATGACGAATTTTCTAGAGATAGATTAAGTGC---	TGGGAAACGTACACAGG	1015
Db	507	TTGACATGATGAGAAATCCCGACAGAGATGTGGAGTGTCTCGACAGAGACCATTAACACAGG	448	
QY	1016	TGCTGCATATGCTGTCTGCAGCTCGTGTCTGTGAGATGTGTGGTTAAGTCCCGCAACGACGC	1075	
Db	447	TGCTGCATATGCTGTCTGTCTGCAGCTCGTGTCTGTGAGATGTGTGGTTAAGTCCCGCAACGAGCG	388	
QY	1076	CAACCTTGTCTAATTAATTTGCCATCATATTTGGTTGGCACTTTAATGAGACTGCGCGTGACA	1135	
Db	387	CAACCTTGTCCATCAGTTGCTACGAAA-----GGGCACTGTGATGGGACTGCGCGTGACA	333	
QY	1136	AACCGGAGGAGATGTGGGGATATACACGTCAAGTCTCTATAGGCCCCTTAATGGGTAGGGCTTACA	1195	
Db	332	AACCGGAGGAGATGTGGGGATATACACGTCAAGTCTCTATAGGCCCCTTAATGGGTAGGGCTTACA	273	
QY	1196	CGTAATACATATGGCGCGCTACAGAGGGTTGCCAACCCCGGAGGGGAGACTAATCTCAGAAA	1255	
Db	272	CGTCAATACATATGGCGCGCTACAGAGGGTTGCCAACCCCGGAGGGGAGACTAATCTCAGAAA	213	
QY	1256	GGCGGTCGTAGTCCGGATCCGAGTCCGAACTCGACTCGACTTCGTGAAGTGGAAATCGCTAGTA	1315	
Db	212	GGCGGTCGTAGTCCGGATCCGAGTCCGAACTCGACTCGACTTCGTGAAGTGGAAATCGCTAGTA	153	
QY	1316	ATCGCGCATACACATGTCTCCGCTGAATACGTTCCCGGGTCTTTGTACACACCGCCCGCTAC	1375	
Db	152	ATCGCGCATACACATGTCTCCGCTGAATACGTTCCCGGGTCTTTGTACACACCGCCCGCTAC	93	
QY	1376	ACCATGGGAGTGGGTTTCAACGAGAGGATAGTCTAACCGTAAGGAGGAGGCGTTGCCAC	1435	
Db	92	ACCATGGGAGGAGGTTTCAACGAGAGGATAGTCTAACCGTAAGGAGGAGGCGTTGCCAC	33	
QY	1436	GGTGAATTCATGACTGGGGTG	1457	
Db	32	GGTGGGCTTCGTGACTGGGGTG	11	
RESULT 13				
US-09-063-898-1				
: Sequence 1, Application US/09063898				
: Patent No. 6319497				
: GENERAL INFORMATION:				
: APPLICANT: Cascida, Jr., Lester E.				
: APPLICANT: Falkinham, III, Joseph Oliver				
: APPLICANT: Cain, Cody C.				
: TITLE OF INVENTION: NON-OBLIGATE PREDATORY BACTERIUM				
: TITLE OF INVENTION: BURKHOLDERIA CASIDAE AND USES THEREOF				
: NUMBER OF SEQUENCES: 1				
: CORRESPONDENCE ADDRESS:				
: ADDRESSEE: Pennie & Edmonds LLP				
: STREET: 1667 K Street, N.W.				
: CITY: Washington				
: STATE: DC				
: COUNTRY: USA				
: ZIP: 20006				
: COMPUTER READABLE FORM:				
: MEDIUM TYPE: Diskette				
: COMPUTER: IBM Compatible				
: OPERATING SYSTEM: DOS				
: SOFTWARE: FASTSEQ Version 2.0				
: CURRENT APPLICATION DATA:				
: APPLICATION NUMBER: US/09/063,898				
: FILING DATE:				
: CLASSIFICATION:				
: PRIOR APPLICATION DATA:				
: APPLICATION NUMBER: 60/044,532				
: FILING DATE: 23-APR-1997				
: ATTORNEY/AGENT INFORMATION:				
: NAME: Mistrock, Leslie				
: REGISTRATION NUMBER: 18,872				
: REFERENCE/DOCKET NUMBER: 8743-006-999				
: TELECOMMUNICATION INFORMATION:				
: TELEPHONE: (212) 7909090				

Query Match	Similarity	Score	ID NO.	Length
Best Local Match	86.4%	1057.6	DB 3	1495
Matches 1262	Conservative 0	Mismatches 184	Indels 14	Gaps 8

  

Query	1	ATTGAGCGCTGCGCGCATGCTTTTCA	CATGCCAGTCGAAAGCGACGCGATGCTTCA	59
Db	3	ATATTAGCGCTGGTGTGACATGCTTTACAGATGCCAAGTCGACGCGCGTGTCTTCA		62
Qy	60	TCGTGTGCGGAGTGTGCGGACGGGTGAGTAATGATCGGAC-GTATCCAGAAAGGCGG		118
Db	63	CTGTGTGCGGAGTGTGCGGACGGGTGAGTAATGATCGGACCAATGCTTGTATGTGCGG		122
Qy	119	TAAAGCATCGAAAGATGTCTTAATACCGCATATCTCTTAAGAGAAAGACAGGAGATCGA		178
Db	123	TAGCCCGCGGAAACCGGATTTAATACCGCATATCGATCTACGAGTGAAGACGGGAGCCTT		182
Qy	179	AAGACCTTGGCGCTTTTGGAGCGGCCGATGTCTGATTAGCTATGTTGTGGGTAAAGGCTT		238
Db	183	CGGCGCTCGCGCTTATGAGGTTGGCCGATGCTGATTAGCTATGTTGTGGGTAAAGGCTT		242
Qy	239	ACCAAGCGCGATCATGTAATGTTGTCTGAGAGACGACACCGACACTGGGACTGAGACA		298
Db	243	ACCAAGCGCGATCATGTAATGTTGTCTGAGAGACGACACCGACACTGGGACTGAGACA		301
Qy	299	CGGCGCGACACTCTTACCGGAGGACGACAGTGGGAAATTTTGGACATAGGCGCGACCTGA		358
Db	302	CGGCGCGACACTCTTACCGGAGGACGACAGTGGGAAATTTTGGACATAGGCGCGACCTGA		361
Qy	359	TCCAGCAATGCGCGTGTGAGTAAGAGGCGCTTGGGTTGTAAAGCTCTTTCACTGAGAA		418
Db	362	TCCAGCAATGCGCGTGTGAGTAAGAGGCGCTTGGGTTGTAAAGCACTTTGTCCGAAA		421
Qy	419	GAAAGGTTACGGTAAATTAATCTGTGACCCCATGACGCTATCCAGAGAGACGCGCTA		478
Db	422	GAAATCTTGTGTTCTAATTAATGACCGCGGAGTACGCTACCGAATAATGACACCGGCTA		481
Qy	479	ACTAGTGCACGACGCGCGGTAATACGTAAGGTCGCAAGCTTAATCGGAATTTACTGGGC		538
Db	482	ACTAGTGCACGACGCGCGGTAATACGTAAGGTCGCAAGCTTAATCGGAATTTACTGGGC		541
Qy	539	GTAAGGAGTGTGCGACGCGCGCTTGTAGTCAAGATGTGAATCCCGGCGTTAACTCGGA		598
Db	542	GTAAGGAGTGTGCGACGCGCGCTTGTAGTCAAGATGTGAATCCCGGCGTTAACTCGGA		601
Qy	599	ATTGCGTTGAAACTCAAGACTTAAGTGTGCGACAGGAGGTGGAATTTCCATGTGTAC		658
Db	602	ACTGATTTGTGAGCTGTGCGACGCTTAAGTGTGCGACAGGAGGTGGAATTTCCATGTGTAC		661
Qy	659	AGTGAATGTGCGAGATTAAGGAACAATGATGAGCGGACGCGCTCTGAGGTAAACA		718
Db	662	AGTGAATGTGCGAGATTAAGGAACAATGATGAGCGGACGCGCTCTGAGGTAAACA		721
Qy	719	CTGACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTTAGATACCTGTGATGTCACG		778
Db	722	CTGACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTTAGATACCTGTGATGTCACG		781
Qy	779	CCCTAAACGATGCACTAGTTGTTGGCGCTTATTAAGCTTGTGTAAAGACTTAACGCT		838
Db	782	CCCTAAACGATGCACTAGTTGTTGGCGCTTATTAAGCTTGTGTAAAGACTTAACGCT		841
Qy	839	GAACTTGAACCGCTGGGAGTACGCTCGAATAATTAACTCAAGGAATTGACGGGAC		898

Db 842 GAAGTTGACCGCTGGGAGTAAGTCGCAAGATTAATACTCAAGAAATTGACGGGAGC 901  
Qy 899 CCGCACAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAAACCTTAACCTACCC 958  
Db 902 CCGCACAAGCGGTGATTAATGATTAATTCGATGCAACCGGAAAACCTTAACCTACCC 961  
Qy 959 TTGACATGTAAGGAATTTCTAGAGTAAGTTAGTCT---TCGGGACGCTTAACGACAG 1015  
Db 962 TTGACATGTAAGGAATTTCTAGAGTAAGTTAGTCT---TCGGGACGCTTAACGACAG 1021  
Qy 1016 TCTGATGATGCTGCTGATGCTGCTGATGATGCTGATGATGCTGATGATGCTGATGATG 1075  
Db 1022 TCTGATGATGCTGCTGATGCTGCTGATGATGCTGATGATGCTGATGATGCTGATGATG 1081  
Qy 1076 CAACCTTGTCAATTAATGCTCAATTTGGTGGGCACTTTAATGAGTCTCCGCTGACA 1135  
Db 1082 CAACCTTGTCTTAATGCTC-----TACGCAAGAGCACTTAAGAGAGACTGCGGTGACA 1136  
Qy 1136 AACCGGAGAAAGTGGGAGTAAGTCAAGTCTCATGGCCCTTAAGGTTAGGGCTTACCA 1195  
Db 1137 AACCGGAGAAAGTGGGAGTAAGTCAAGTCTCATGGCCCTTAAGGTTAGGGCTTACCA 1195  
Qy 1196 CTTAATACATGCGCGGTACAGAGGTTGCAACCGCGAGGGGAGCTAATCTCAGAAA 1255  
Db 1196 CTTAATACATGCGCGGTACAGAGGTTGCAACCGCGAGGGGAGCTAATCTCAGAAA 1254  
Qy 1256 GCGCGTGTAGTCCGATGCGAGTCTGCAACTGACTCCGTTGAATGCGAATGCTAGTA 1315  
Db 1255 ACCGATGTAGTCCGATGCGAGTCTGCAACTGACTCCGTTGAATGCGAATGCTAGTA 1314  
Qy 1316 ATCGCGGATCAGATGCGCGGTAAGTAAGTCCCGGCTTGTACACACCGCCGCTAC 1375  
Db 1315 ATCGCGGATCAGATGCGCGGTAAGTAAGTCCCGGCTTGTACACACCGCCGCTAC 1374  
Qy 1376 ACCATGAGAGTGGGTTTCAACGAGAGAGTACTTAACCGTAAGG-AGGGCGCTTGCCA 1434  
Db 1375 ACCATGAGAGTGGGTTTCAACGAGAGAGTACTTAACCGTAAGG-AGGGCGCTTGCCA 1434  
Qy 1435 CCGTAGATTCACTGCTGGG 1454  
Db 1435 CCGTAGATTCACTGCTGGG 1454

RESULT 14  
US-09-985-846-1  
Sequence 1, Application US/09985846  
Patent No. 6689357  
GENERAL INFORMATION:  
APPLICANT: Casida, Jr., Lester E.  
Falkinham, III, Joseph Oliver  
Cain, Cody C.  
TITLE OF INVENTION: NON-OBLIGATE PREDATORY BACTERIUM  
BURKHOLDERIA CASIDAE AND USES THEREOF  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1667 K Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/985,846  
FILING DATE: 06-No. 6689357-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/063,898  
FILING DATE: 22-APR-1998

APPLICATION NUMBER: 60/044,532  
FILING DATE: 23-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mastrock, Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 8743-006-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)7909090  
TELEFAX: (212)8699741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Genomic DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-985-846-1  
Query Match 72.6%; Score 1057.6; DB 3; Length 1495;  
Best Local Similarity 86.4%; Pred. No. 0;  
Matches 1262; Conservative 0; Mismatches 184; Indels 14; Gaps 8;  
Qy 1 ATTTGAACGCTGGCGCATGCTTTTACA-CATGCAAGTCGAACGGCAGACGATGCTTGA 59  
Db 3 ATTTAAGCTGTTTGATGCTTTTACAAGATGCAAGTCGAACGGCAGACGATGCTTGA 62  
Qy 60 TTTGTGCGAGTGGCGGAGCGGTTGATGATGCTGCAAC-GTATCCGAAGAGGGGG 118  
Db 63 CTTGTGCGAGTGGCGGAGCGGTTGATGATGCTGCAACATGCTGCTGTGTGGGGGA 122  
Qy 119 TAAAGCATGGAAGATGCTAATACCGCATTAATCTTAAGAGGAAAGCAGGGATGCA 178  
Db 123 TAAAGCATGGAAGATGCTAATACCGCATTAATCTTAAGAGGAAAGCAGGGATGCA 182  
Qy 179 AAGACCTTGGCTTTTGGAGCGCGCATGCTGATTAAGTGGTGGGTTAAAGGCTT 238  
Db 183 AAGACCTTGGCTTTTGGAGCGCGCATGCTGATTAAGTGGTGGGTTAAAGGCTT 242  
Qy 239 ACCAAGGCGACATCATGATGTTGTTGAGAGAGCAGACCACTGGGACTGAGACA 298  
Db 243 ACCAAGGCGACATCATGATGTTGTTGAGAGAGCAGACCACTGGGACTGAGACA 301  
Qy 299 CCGCCGAGCTCTACGGGAGGAGCAGAGGGGAAATTTTGAACAATGGCGCAGGCTGA 358  
Db 302 CCGCCGAGCTCTACGGGAGGAGCAGAGGGGAAATTTTGAACAATGGCGCAGGCTGA 361  
Qy 359 TCCAGCAATGCGCGTGAAGTGAAGAGGCTTCCGGTTGTAAAGCTTTTCAGTCGAGA 418  
Db 362 TCCAGCAATGCGCGTGAAGTGAAGAGGCTTCCGGTTGTAAAGCTTTTCAGTCGAGA 421  
Qy 419 GAAAAGTTACGGTAATTAATGTTGATGCCATGACGCTATGACAGAAAGACCGGCTA 478  
Db 422 GAAATCTTGTGTTCTAATATAGCCGGGGATGACCGGTACCGAAGATTAAGACCGGCTA 481  
Qy 479 ACTACGTCGACAGCGCGGTTAATGTAAGGTTGCAACGTTAATCGAATTAAGTGGG 538  
Db 482 ACTACGTCGACAGCGCGGTTAATGTAAGGTTGCAACGTTAATCGAATTAAGTGGG 541  
Qy 539 GTPAAGGTTGCGCAGCGGCTTGTAAAGTGTGACAGATGTAATCCCGGCTTAACCTGGA 598  
Db 542 GTPAAGGTTGCGCAGCGGCTTGTGTAAAGTGTGACAGATGTAATCCCGGCTTAACCTGGA 601  
Qy 599 ATTGCTTTGAACCTAACAAGTGAAGTGTGACAGAGGAGTGTGAATTCATGTATAC 658  
Db 602 ACTGATGTTGATGCTGACAGGCTGATGATGACAGAGGAGTGTGAATTCATGTATAC 661  
Qy 659 AGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 718  
Db 662 AGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721  
Qy 719 CTGACGCTATGACGAAAGCGTGGGAGCAAAACAGATTAGTACCTGTGTAGTCAAG 778

```

Db      722  CTGACGCTCATGCAACGAAAGCGTGGGAGCAACAGATTAAGTACCTGTGTGTCAAG  781
QY      779  CCCTAAACGATGTCACATAGTGTGGGCTTATTAAGCTTGTGAACGAAGCTAACGCGT  838
Db      782  CCTTAACGATGTCACATAGTGTGGGGAATTCTATTTCTTAGTAAGGTAGCTAACGCGT  841
QY      839  GAAGTTGACCGCTCGGGGAGTAGTCGCTGGCAAGATTAAACTCAAAAGGAATTGACGGGGAC  898
Db      842  GAAGTTGACCGCTCGGGGAGTAGTCGCTGGCAAGATTAAACTCAAAAGGAATTGACGGGGAC  901
QY      899  CCGCAAGAGCGGTGATTTATGTGATTAATTTCGATGCAACGCGCAAAAACCTTACTTACC  958
Db      902  CCGCAAGAGCGGTGATTTATGTGATTAATTTCGATGCAACGCGCAAAAACCTTACTTACC  961
QY      959  TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTCT--TCGGAAACGCTAACACAGG  1015
Db      962  TTGACATGTGTCGAATCCCGCTGAGAGGTGGGATGTCTCGAAAGAGAACCGGCGACAGG  1021
QY      1016  TGTCTGATGCGTGTGCTGTCAGTCTGTGTCGTGAATGTTGGGTTAAATGCCGCAACGAGC  1075
Db      1022  TGTCTGATGCGTGTGCTGTCAGTCTGTGTCGTGAATGTTGGGTTAAATGCCGCAACGAGC  1081
QY      1076  CAACCTTGTGATTAATTGCGCATCATTTGTTGGGCACTTAAATGAGACTGCCGCTGACA  1135
Db      1082  CAACCTTGTCTTAGTTGC-----TACGCAAGAGCACTCTAAGAGACTGCCGCTGACA  1136
QY      1136  AACCGAGGAAGTGGGGATGACGTCAAGTCTCATGCGCTTATGGGTAGGGCTTGACA  1195
Db      1137  AACCGAGGAAGTGGGGATGACGTCAAGTCTCATGCGCTTATGGGTAGGGC-TGACA  1195
QY      1196  CGTAAATCAATGGCGCGTGAACAAGGGTGTCCAAACCCGCGAGGGGAGCTAATCTTCAGAA  1255
Db      1196  CGTCAATCAAAATGGTTCGGAACAGAGGGTGTCC-AACCCGGAAGGGGAGCTAATCCAGAA  1254
QY      1256  GCGCGGTGATCCGGGATCGGATGTCGAACTGCACTCGACTCGTGAAGTCCGGAATCGCTAGTA  1315
Db      1255  ACCGATGCTAATCCGGGATTCGACTCTGCACTCGAGTCAATGAACTGGAAATCGTAGTA  1314
QY      1316  ATCCGCGATCAGCATGTCGCGGTGATACGTTCCGGGTCCTTGTACACACCGCCCGTCAAC  1375
Db      1315  ATCCGCGATCAGCATGTCGCGGTGATACGTTCCGGGTTTGTATACACCGCCCGTCAAC  1374
QY      1376  ACCATGGAGTGGGTTTTCACCAAGACGAGTAGTCTAAACGCTAAGG-AGGGCGCTTGCCA  1434
Db      1375  ACCATGGGAGTGGGTTTTCACCAAGAGTGGTAGTCTAACCGCAAGGAAGAACGCTGCCCA  1434
QY      1435  CGGTGAGATTCAATGACTGGG  1454
Db      1435  CGGTGAGATTCAATGACTGGG  1454

RESULT 15
US-09-228-184-1
; Sequence 1, Application US/09228184
; Patent No. 6322782
; GENERAL INFORMATION:
; APPLICANT: WALKER, Harrell L.
; APPLICANT: HIGGINBOTHAM, Lawrence R.
; TITLE OF INVENTION: CONTROL OF CYANOBACTERIA WITH A BACTERIUM
; FILE REFERENCE: 013243-0007
; CURRENT APPLICATION NUMBER: US/09/228,184
; CURRENT FILING DATE: 1999-01-11
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1540
; TYPE: DNA
; ORGANISM: Unknown Organism
FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Bacterium SG-3;
; OTHER INFORMATION: gram negative; rod-shaped; exhibits flagellate
; OTHER INFORMATION: motility; pathogenic to cyanobacteria and algae;

```

OTHER INFORMATION: yellow colonies on BG-11 medium suppl. with tryptic  
US-09-228-184-1

Query Match	71.9%;	Score 1048.2;	DB 3;	Length 1540;
Best Local Similarity	84.5%;	Pred. No. 0;		
Matches 1239;	Conservative	0;	Mismatches 218;	Indels 10;
				Gaps 5;

Oy	1	ATTGAAAGCGTGGCGGCAATGCTTTAACAATGCAATGCAACGGGAGCAGGATCTTGAT	60
Db	24	AGTAAAGCGTGGCGGCGCTTAACATCATCAAGTGGAAACGGCAGCAGCACTAGCAAT	83
Oy	61	CT----GGTGGCAGTGGCGGACGGGTGAATTCATCGGAACTGATCCGAAGAGGGG	116
Db	84	ACTGTGGGTGGCGAGTGGCGGACGGGTGAAGAAATACATCGGAATCTGGCCAGTCTGGGGG	143
Oy	117	GGTAAAGCATGCAAAAGATGTGCTAATACCCCATATATCTTAAGAGGAAGACAGGGGATC	176
Db	144	GATTAACATTAAGAAACTTAATGCTTAATACCGATACGACCTTAACGGGTGAAGCGGGGATC	203
Oy	177	GAAAGACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTTAAAGGC	236
Db	204	GCAAGACCTGCGGATTTGGATGGATGAGCGGATGTCCGATTACCTAGTTGGCGGGGTAAATGGC	263
Oy	237	CTAACCAAGCGACGATCAATGATTTGGTCTGAAGAGACGACACGACACACTGGGACCTGAGA	296
Db	264	CCACCAAGCGCACATCGTAGCTGTGCTGAAGAGATGATCAGCACACTGGAATCTGAGA	323
Oy	297	CACGGCCCAACTCTTACGGGAGGACACAGTGGGAAATTTTGGACAATGGGCGCAAGCCT	356
Db	324	CACGGTCCAGCTCTTAACGGGAGGACACAGTGGGAAATTTTGGACAATGGGCGCAAGCCT	383
Oy	357	GATTCAGCAATGCCGCGTAGTGAAGAAAGCCTTCGGGTTGTAAAGCTCTTTCAGTGAG	416
Db	384	GATCCAGCCATGCCGCGGTGTGAAGAGGCTTCGGGTTGTAAAGCCTTTCAGTGAG	443
Oy	417	AAGAAAGGTTAAGCTTAATTAATGTAACCATGACCGTATTCGACAGAGAAAGCACCGGC	476
Db	444	AAAGAAAGCATGATTAATTAATCTTCGGTGTCTGACCGTATTCGAGAAATTAAGCACCGGC	503
Oy	477	TAACTACGTGCACGACGCCGCGGTAAATTCGTAAGGTCAGAGCTTAATTCGGAATTAATCG	536
Db	504	TAACTTCGTGCACGACGCCGCGGTAAATTCGTAAGGTCAGAGCTTAATTCGGAATTAATCG	563
Oy	537	GCGTAAAGGATGCGGAGCGGCGCTTGTAAAGCAGATGTAATCCCGGGCTTAACCTGG	596
Db	564	GCGTAAAGGATGCGGAGCGGCGCTTGTAAAGCAGATGTAATCCCGGGCTTAACCTGG	623
Oy	597	GAAATTCGCTTGAACCTAACAGCTAGAGTGTGCAAGGAGGAGGTGAATTCATGTGTA	656
Db	624	GAAATTCGCTTGAACCTAACAGCTAGAGTGTGCAAGGAGGAGGTGAATTCATGTGTA	683
Oy	657	GCAATGGAATGCTTAAGATTAATGAAGAAATTCATGATGCGAAGGACGCTCTGAGGTTAA	716
Db	684	GCAATGGAATGCTTAAGATTAATGAAGAAATTCATGATGCGAAGGACGCTCTGAGGTTAA	743
Oy	717	CAGTACGCTCATGCAAGAAACGCTGGGAGCAACAGATTAGATCCTGTGATGTCA	776
Db	744	CAGTACGCTCATGCAAGAAACGCTGGGAGCAACAGATTAGATCCTGTGATGTCA	803
Oy	777	GCGCTTAACCATGTCACTAGTGTGGGCTTAATAGG--CTTGGTAAAGAGCTAAC	834
Db	804	GCGCTTAACCATGTCACTAGTGTGGGCTTAATAGG--CTTGGTAAAGAGCTAAC	863
Oy	835	GCGTAAAGTTGACCGCTGGGAGTACCGGTGCAAGATTAAACTCAAGGAATTTGACGG	894
Db	864	GCGTAAAGTTGACCGCTGGGAGTACCGGTGCAAGATTAAACTCAAGGAATTTGACGG	923
Oy	895	GGAACCGGACAAAGCGGTGATTAATGTGATTAATTCGATGGAAGCGGAAAACTTTACT	954
Db	924	GGAACCGGACAAAGCGGTGATTAATGTGATTAATTCGATGGAAGCGGAAAACTTTACT	983
Oy	955	ACCCTTGACATGACGAATTTTCTAGAGATGATTAAGTGTG--CTTCGGGAAGCTTAACCA	1013

Db	964	GGCCTTGACATGCACGGAACTTTCCAGAGATGGATTGGTGGCTTCGGGAACCGGTGACACA	1043
Qy	1014	GGTGTGCATGGCTGTGCTCAGCTCGTGTGCTGAGATGTTG3GTTTAAGTCCCGCAACGAG	1073
Db	1044	GGTGTGCATGGCTGTGCTCAGCTGTGTCTGGAGATGTTGG3TTTAAATCCCGCAACGAG	1103
Qy	1074	GGCAACCCCTTGCTAATTAATGCGATCA--TTTGGTGG3CACTTAAATGACATCGCGGT	1131
Db	1104	GGCAACCCCTTGCTAATTAATGCGATCAATATGTTGG3AACTCTHAGGAAACCGCGCGC	1163
Qy	1132	GACAAACCGAGAGAGTGGGGATGACGTCAATCTCTATG3CCCTTATGSGTAAAGGCTT	1191
Db	1164	GACAAACCGAGAGAGTGGGGATGACGTCAAGTCAATCATATG3CCCTTATGSGTAAAGGCTT	1223
Qy	1192	CACAGTAAATCAATGGGCGGTACAGAGGGTTGGCAACCCGCGAGGGGAGACTAATCTCA	1251
Db	1224	CACAGTAAATCAATGGGCGGTACAGAGGGTTGGCAACTCCGAGAGTGAAGCAATCCCA	1283
Qy	1252	GAAACCGGCTCGTAGTCCGGATCGGAGTCTGCACATCTCGGTGAAGTGGAAATGCT	1311
Db	1284	GAAACCCCATCTCACTCCGGATTTGGAAGTCTGCAATCTCAATGAAGTGGAAATGCT	1343
Qy	1312	AGTATATCGCGGATCAGCA-TGTGCGGGGTGAATACCTTCCCGGGCTTGTACACACCGCC	1370
Db	1344	AGTATATCGCAATCAGCAATGTCTGGGTGAATACCTTCCCGGGCTTGTACACACCGCC	1403
Qy	1371	GTCAACACATGGAGATGG3TTTCAACAAGACAGTACTTACACGTAAAGAGGG3CGTT	1430
Db	1404	GTCAACACATGGAGATTTTGTGGACACAAGAGAGTACTTAAACCGAAGGGGG3CGTT	1463
Qy	1431	GCCACGCTGAAGTTATGACTGGGGTG	1457
Db	1464	GCCACGCTGTGAGCCGATGACTGGGGTG	1490

```

RESULT 16
US-09-967-376-1
; Sequence 1, Application US/09967376
; Patent No. 6482635
; GENERAL INFORMATION:
; APPLICANT: WALKER, Harrell L.
; APPLICANT: HIGGINBOTHAM, Lawrence R.
; TITLE OF INVENTION: CONTROL OF CYANOBACTERIA WITH A BACTERIUM
; FILE REFERENCE: 013243-0007
; CURRENT APPLICATION NUMBER: US/09/967.376
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1540
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Bacterium SG-3; gram negative;
; OTHER INFORMATION: rod-shaped; exhibits flagellate motility; pathogenic to
; OTHER INFORMATION: cyanobacteria and algae; yellow colonies on BG-11 medium suppl.
US-09-967-376-1

Query Match          71.9%; Score 1048.2; DB 3; Length 1540;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1239; Conservative 0; Mismatches 218; Indels 10; Gaps 5;

QY      1  ATTGAACGCTGGCGGCGCATCTTTACATCATGCAAGTCGAACGGCAGCAGCATGCTTGCAAT 60
DB      24  AGTGAACGCTGGCGGCGGCGCATCTTACATCATGCAAGTCGAACGGCAGCAGCATGTAACAT 83
QY      61  CT----GGTGGCGAGATGGCGGCGGGGTGATGTAATGCATCGGAAAGTATCCAGAAAGGGG 116
DB      84  ACTGGGGGTGGCGAGTGGCGGAGCGGGTGAAGAAATCATCGGAATTCGCGCATGCTGTGGG 143
QY      117  GGTAAAGCATGAAAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGACAGGGATC 176

```

Db	1144	GATTAACATAGGAAACTTATGCTAATACCGCATACGACCTACGGGTGAAGCGGGGATC	203
Qy	1177	GAAAGACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAAGTAAGTTGGTGGGTAAAGC	236
Db	204	GCAAGACCTCGCGCATTTGGATGAGCCGATGTCCTATTAGCTAGTTGGCGGGTAAATGGC	263
Qy	2237	CTACCAAGGCAACGATTCAGTAGTTGGTGTCTGAGAGGACGACCACTGGGACTGAGA	296
Db	264	CCACCAAGGCAACGATTCGATGCTGGTCTGAGAGGATGATACGCCACCTGGAACTTGAGA	323
Qy	297	CACGCGCCGACACTCCCTACGGGAGGACGAGTGGGAAATTTTGGACAAATGGCGCAAGCT	356
Db	324	CACGGTCCAGACTCCTACGGGAGGACGAGTGGGAAATATTGGACAAATGGGCGCAAGCT	383
Qy	357	GATCCAGCAATGCGCGGTGAGTGAAGAAAGGCGCTTGGGTGTAAAGCTCTTCACTCGAG	416
Db	384	GATCCAGCAATGCGCGGTGAGTGAAGAAAGGCGCTTGGGTGTAAAGCGCTTTGTGCCGA	443
Qy	417	AAGAAAGGTTACGGTAAATTAATCGTACCCATGACGGTATTCGACAGAAAGCACCGGC	476
Db	444	AAGAAAGGTTACGGTAAATTAATCGTACCGTATTCGACGGTATTCGACAGAAAGCACCGGC	503
Qy	477	TAACTACGTGCGACGAGCGCGGCTAATACGTAGGGTGCAGCGTAAATCGAATTAATCTGG	536
Db	504	TAACTCTGCGCGACGAGCGCGGCTAATACGAAGGGTGCAGCGTAAATCGAATTAATCTGG	563
Qy	537	GCGTAAAGGGTGCAGGCGCGCTTTGTAAGTCAAGTGTGAATCCCGGGCTTAACTTGG	596
Db	564	GCGTAAAGGGTGCAGGCGCGCTTTGTAAGTCAAGTGTGAATCCCGGGCTTAACTTGG	623
Qy	597	GAAATGCGCTTTGAAACTAACAAGCTAGTGTGCGACAGGGAGGTGAATTCATGTTGA	656
Db	624	GAAATGCGCTTTGAAACTAACAAGCTAGTGTGCGACAGGGAGGTGAATTCATGTTGA	683
Qy	657	GCAGTGAATGCGTGAAGATATGGAAGAAACAATCGATGCGAAGGACGCTCCTGGGTAA	716
Db	684	GCAGTGAATGCGTGAAGATATGGAAGAAACAATCTGTGGCGAAGGCGGCATCTGGACCA	743
Qy	717	CACCTGACGCTATGACGCAAGAAACGCTGGGAGCAACAGAGTATTAATCCTGTGATGCCA	776
Db	744	CACCTGACGCTATGAGGACGAAAGCGTGGGAGCAACAGAGTATTAATCCTGTGATGCCA	803
Qy	777	CGCCTTAAACGATGTCACTAGTTGTTGGGCTTATTAGG--CTTGTGTAAGAAAGCTAAC	834
Db	804	CGCCTTAAACGATGTGGAATCGAATGTTGGGTGCAACTAGGCACTCAGTATGCAAGCTAAC	863
Qy	835	GCGTGAATGTAACGCGCTGGGGAGTACCGGTGCGCAAGTTAAATCTCAAGGAATTGACGG	894
Db	864	GCGTGAATGTTGCGCGCTGGGGAGTACCGGTGCGCAAGTTAAATCTCAAGGAATTGACGG	923
Qy	895	GGACCCGACAAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAAACTTTACT	954
Db	924	GGGCCCCGACAAAGCGGTGAGATATGTGTTTAATTCGATGCAACGCGAAAACTTTACT	983
Qy	955	AACCTTGACATGTAGCGAATTTTCTAGAGATAGATTAGTG-CTTGGGAAAGCTTAACCA	1011
Db	984	GGCCTTGACATGTAGCGAATTTTCTAGAGATAGATTAGTGCTTGGGAAACCGTGAACCA	1043
Qy	1014	GGTGTGATAGGCTGTGCTGACAGCTGTGTCGTAAGTGTGGGTTAAAGTCCCGCAACGAG	1072
Db	1044	GGTGTGATAGGCTGTGCTGACAGCTGTGTCGTAAGTGTGGGTTAAAGTCCCGCAACGAG	1100
Qy	1074	CGCAACCTTGTCAATTAATGCGCAATCA--TTTGTTGGGCACTTTAATGAGACTGCGGT	1133
Db	1104	CGCAACCTTGTCTTACTTAATGCGCAACGTAATGTTGGGAACTCTAAGGGAAGACGCGCGG	1166
Qy	1132	GACAAACCGAGAGAAAGTGGGATGACGTCAAGTCTCTATGGCCTTTATGGGTAGGCTT	1199
Db	1164	GACAAAGCGGAGAGAAAGTGGGATGACGTCAAGTCTATCATGAGCCTTTACGCGCAAGGCTTA	1222
Qy	1192	CACACGTAAATCAATGTCGGCTGACAGAGGTTGCCAACCCGCGAGGGGAGGTAATCTCA	1255
Db	1224	CACACGTAAATCAATGTCGGGACAGAGAGGCTGCAAACTTCGCGAAGTGAAGCAATCTCCA	1288





```

RESULT 18
US-09-735-567-3
; Sequence 3, Application US/09735567
; Patent No. 6608190
; GENERAL INFORMATION:
; APPLICANT: Brumcci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BC1033 US NA
; CURRENT APPLICATION NUMBER: US/09/735,567
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Denitrifying bacteria
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RA1
US-09-735-567-3

```

```

Query Match      70.2%; Score 1023.2; DB 3; Length 1450;
Best Local Similarity 84.3%; Freq. No. 0;
Matches 1215; Conservative 0; Mismatches 213; Indels 14; Gaps 5;

QY 1 ATGAAGCTGCGCGCATGCTTACACATGCAAGTCGAAACGCGACGCGATCTTGCAAT 60
DB 3 ATTGAACGCTGGCGGAATGCTTTACACATGCAAGTCGAAACGCGCGCGGCA----AC 58
QY 61 CTGGTGGCGAGTGGCGGACGCGATGATGATCGGAACTGATCCGAAAGAGCGGCGTA 120
DB 59 CTGGCGCGGAGCGCGCAACGCGTGAATGATACATCGGAACGTGCCAGAGTGGGGAATA 118
QY 121 AGCGATCGAAAGATGCTGCTAATCCGCTATCTCTTAAGAGAGAAAGCGGGATGAAA 180
DB 119 GCCCGCGGAAAGCGGATTAATACCGCATGTGATCTGAGGATAAAGTGGGGAACGCA 178
QY 181 GACCTGCGCTTTTGGAGCGCGCATGCTGATTAAGTGGTGGGCTAAAGCGCTAC 240
DB 179 GCGCTACCGCTTTGAGAGCGCGCATGCGAATTAAGTGGTGGGCTAAAGCGCTAC 238
QY 241 CAAGCGACGATGATGATGCTGATGAGAGACGACCACTGGGACTGAGACAG 300
DB 239 CAAGCTGCGATGATGATGCTGATGAGAGATGATGACCACTGGGACTGAGACAG 298
QY 301 GCCCAACTCTTACGCGGAGCGACAGATGGGGAATTTTGAACATGGGCGCAACCTGATC 360
DB 299 GCCCAACTCTTACGCGGAGCGACAGATGGGGAATTTTGAACATGGGCGCAACCTGATC 358
QY 361 CAGCAATGCGCGCTGATGAGAAAGCGCTTGGGTTGTAAGCTCTTTCAGTGAAGA 420
DB 359 CAGCAATGCGCGCTGATGAGAAAGCGCTTGGGTTGTAAGCTCTTTCAGTGAAGA 418
QY 421 AAGGTTACGTTAATTAATCTGATGACCCATGACGCTATGACAGAAAGACGCGCTAC 480
DB 419 AAGGCTCTCTTAATATACAGGGGACATATGACGCTACCGTAAGTAAGACGCGCTAC 478
QY 481 TAGTCCGACAGACCGCGCTAATACGTAAGGTGCAACCGTTATCCGAATTAATCTGCGCT 540
DB 479 TAGTCCGACAGACCGCGCTAATACGTAAGGTGCAACCGTTATCCGAATTAATCTGCGCT 538
QY 541 AAGGTTGCGACGCGCGCTTGAAGTCAATGTAATCCCGGCTTAACTGGGAAT 600
DB 539 AAGGTTGCGACGCGCGCTTGAAGTCAATGTAATCCCGGCTTAACTGGGAAT 598
QY 601 TGGTTTGAATCAAAAGCTAGATGTGCGACAGGAGGTGAATTCATGTGTACAG 660
DB 599 TGGCTTTGATGACAAAGCTGAGTGGCGGACAGAGGGGATGTAATTCGCGGTGACAG 658

```

```

QY 661 TGAATGCTAGATATGAGAAACATGATGCGAAAGGACGCTCTCGGTTAACT 720
DB 659 TGAATGCTAGATATGCGAGAAACACGATGTTGAAGGCAATCCCGTGGCATGACT 718
QY 721 GACGCTATGACGAAAGCGTGGGAGCAACAGATTAATACCTGTTAGTCCAGCC 780
DB 719 GACGCTATGACGAAAGCGTGGGAGCAACAGATTAATACCTGTTAGTCCAGCC 778
QY 781 CT-AAAAGTTCACACTAGTTTGTGGGCTTTTATAGGCTTGTGTAACGAAGCTACGCG 839
DB 779 CTAAAAAGTTCACACTAGTTTGTGGGATTAATTTCTCACTAACGAAGCTACGCG 838
QY 840 AAGTTGACCGCTGGGAGATGAGTTCGCAAGATTAATCAAGGATTAAGCGGAGC 899
DB 839 AAGTTGACCGCTGGGAGATGAGTTCGCGCAAGGATTAATCAAGGATTAAGCGGAGC 898
QY 900 CGCAGAGCGGTGATGATGATGATTAATGATGATGATGATGATGATGATGATGAT 959
DB 899 CGCAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
QY 960 TGACATGATGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGAT 1016
DB 959 TGACATGATGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGAT 1018
QY 1017 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
DB 1019 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078
QY 1077 AACCTTGTGATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1136
DB 1079 AACCTTGTGATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1133
QY 1137 ACCGAGAGAGGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196
DB 1134 ACCGAGAGAGGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
QY 1197 GTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256
DB 1194 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
QY 1257 CCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1316
DB 1254 CCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
QY 1317 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375
DB 1314 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373
QY 1376 ACCATGAGAGTGGGTTTCAACAGAGATGATGATGATGATGATGATGATGATGAT 1435
DB 1374 ACACATGAGAGGTTTCAACAGAGATGATGATGATGATGATGATGATGATGAT 1433
QY 1436 GG 1437
DB 1434 GG 1435

```

```

RESULT 19
US-08-632-470-53
; Sequence 53, Application US/08632470
; Patent No. 5976791
; GENERAL INFORMATION:
; APPLICANT: MABIAT, CLAUDE
; TITLE OF INVENTION: NUCLEOTIDE FRAGMENTS CAPABLE OF
; FILE OF INVENTION: HYBRIDIZING SPECIFICALLY TO RICKETTSIA DNA OR RNA AND
; TITLE OF INVENTION: THEIR USE AS PROBES OR PRIMERS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSER: OLIF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA

```



COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/632,470  
FILING DATE: 08-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BERRIDGE, WILLIAM P  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: MPB 38238  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)-836-6400  
TELEFAX: (703)-836-2787  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1484 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-632-470-53

Query Match 70.2%; Score 1022.4; DB 2; Length 1484;

Best Local Similarity 82.9%; Pred. No. 0; Mismatches 242; Indels 7; Gaps 4;

Matches 1208; Conservative 0;

1 ATTGAACGCTGGCGGCAATGCTTTACACATGCAAGTCGAAACCGGACGAC--GGATGCTTGC 58  
28 ATTGAACGCTGGCGGCAATGCTTTACACATGCAAGTCGAAACCGGACGACGAGGAGTTC 87  
59 ATCTGTGGC--GAGTGGCGGACGGGTGAGTAATGATGATGCAAGTATCCAGAAAGGGG 116  
88 TCCTGGCGGCGAGAGTGGCGGACGGGTGAGTAATGATGATGCAAGTATCCAGAAAGGGG 147  
117 GGTACGATGCAAGAGTGTCTAATCCGATATCTCTAAGAGGAAAGGAGGAGGATC 176  
148 GATTAACCTGGGAAATCTGGGCTAATACCGTATATCTTTGGGCAAGGAGGAGATC 207  
177 GAAAGACCTTGGCTTTGGAGCGGCGGATGCTGATTAAGTATGCTGGGTGAGTAAAGGC 236  
208 TTCCGAGCTCGTGTCTAATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 267  
237 CTACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 296  
268 CTACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327  
297 CACGGCCGACGATCTCTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 356  
328 CACGGCCGACGATCTCTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 387  
357 GATCAGCAATGCGCGGTGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 416  
388 GATCAGCAATGCGCGGTGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 447  
417 AAGAAAGGTTACGGTAAATATGCTGACCCATGACCGGTATGACAGAAAGAGACCGGC 476  
448 AAGAAAGTCTCAAGGTTAATATCTGAGGCGTTGACGTTGACCAAGAAAGAGACCTGC 507  
477 TAACTAGTGGCGACGAGCGCGGTATATAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 536  
508 TAACTAGTGGCGACGAGCGCGGTATATAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 567  
537 GCGTAAAGGAGTGGCGGCGGTATATAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 596  
568 GCGTAAAGGAGTGGCGGCGGTATATAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 627  
597 GAATTCGTTTAAATCAAGAGTGAAGTGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 656

628 GAATTCGACCCGATATCTGAGTATCTTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 687  
657 GCAAGTAAATGCTGATGATATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 716  
688 GCGGTAAATGCTGATGATATCGAAAGAGACAGAGTGGCGAAAGGAGGAGGAGGAGGAGGAGGAG 747  
717 CACTGACGCTCATGACGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 776  
748 TACTGACACTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 807  
777 CGCCCTAAACGATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 836  
808 CGCGTAAAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867  
837 GTGAAGTGAACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 896  
868 GTTAAGTTCCTCCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 927  
897 ACCCGCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 956  
928 GCGCGCAAGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 987  
957 CTTGACATGATGAGGATTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1015  
988 CTTGACATGATGAGGATTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1047  
1016 TGCTGAT 1075  
1048 TGCTGAT 1107  
1076 CAACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133  
1108 CAACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167  
1134 CAACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193  
1168 TAAACCGAGAGGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1227  
1194 CACGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253  
1228 CACGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1287  
1254 AAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313  
1288 AACTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347  
1314 TAACTGCGGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373  
1348 TAACTGCGGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1407  
1374 ACACATGAGGAGTGGGTTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1433  
1408 ACACATGAGGAGTGGGTTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1467  
1434 ACGGTGAT 1450  
1468 ACGGTGAT 1484

RESULT 20  
US-09-793-920A-1  
Sequence 1, Application US/09793920A  
Patent No. 6479621  
GENERAL INFORMATION:  
APPLICANT: Canon Inc.  
TITLE OF INVENTION: Polyhydroxyalkanoate containing 3-hydroxyphenylalkanoic acid  
FILE REFERENCE: 4396021  
CURRENT APPLICATION NUMBER: US/09/793, 920A  
CURRENT FILING DATE: 2001-02-28  
NUMBER OF SEQ ID NOS: 1  
SEQ ID NO 1  
LENGTH: 1501

```
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii 161 strain.
US-09-793-920A-1

Query Match      70.1%; Score 1021; DB 3; Length 1501;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 235; Indels 6; Gaps 5;

QY      3 TGAACGCTGGCGGCATGCTTTTACATGCAATGCAATGCGACAGCGAGTGCCTTGCATCT 62
DB      1 TGAACGCTGGCGGCATGCTTTTACATGCAATGCAATGCGACAGCGAGTGCCTTGCATCT 59
QY      63 GGTGGCGAGTGGGCGACGGGTGAGTATGCAATGCGAACGATTCAGAAAGAGGGGGGTAC 122
DB      60 GAAATTC-AGCGGCGAGACGGGTGAGTATGCAATGCGAACGATTCGCTGGTGGTGGGAGAAC 118
QY      123 GCATCGAAAGATGTGCTAATATCCGATATATCTTAAGSAGAAAGCAGGGAGTCGAAGA 182
DB      119 GTCTCGAAAGGAGCGCTAATATCCGATATGCTTAAGSAGAAAGCAGGGAGTCCTCGGG 178
QY      183 CCTTGGCGCTTTTGAAGCGCGCATGCTCTGATTAAGTCTGATGCTGGGTAAAGGCTTACCA 242
DB      179 CCTTGGCGCTATCAGATGAGCTTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 238
QY      243 AGGCGACGATCAGTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 302
DB      239 AGGCGACGATCAGTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 298
QY      303 CCAGACTCTCTACCGGAGGAGCAGCAGTGGGGAATTTTGGACAAATGGGCGCAGCCTGATCCA 362
DB      299 CCAGACTCTCTACCGGAGGAGCAGCAGTGGGGAATTTTGGACAAATGGGCGCAGCCTGATCCA 358
QY      363 GCAATGCCCGGTGATGTAAGAAAGGCTTGGGTTTGAAGCTCTTTCAGTCGAGAAAGAA 422
DB      359 GCCATGCCCGGTGATGTAAGAAAGGCTTGGGTTTGAAGCTCTTTCAGTCGAGAAAGAA 418
QY      423 AGGTTACGGTAATTAATTCGATCCCATGACGATGACAGAAAGAACACCGGCTAATCTA 482
DB      419 GGCATTAACCTTAATTAATTCGATCCCATGATGTTTGAAGCTTAACGAAATTAAGCACCAGGCTAATCTC 478
QY      483 CGTGCCAGCAGCCGCGTAAATACGTAAGGTGCAAGCGTTAATCGGAATTAATCTGGGCGTAA 542
DB      479 TGTGCCAGCAGCCGCGTAAATACGTAAGGTGCAAGCGTTAATCGGAATTAATCTGGGCGTAA 538
QY      543 AGGCGTGGCAGCGGCTTGAAGTCAATGTAAGTCAATCCCGGCTTAATACCTGGGAATGG 602
DB      539 AGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 598
QY      603 CGTTTGAATTAACAAAGCTAGAGTGGGAGGAGGAGTGGAAATTCATGATGTAAGCAGTG 662
DB      599 CATTCAAACTGACAAAGCTAGAGTGGGAGGAGGAGTGGAAATTCCTGTGTAGCGGTG 658
QY      663 AAATGCGTAGAGTATGGAAGAACATGATGCGAAGGAGCCTCTGGGTTAACATGTA 722
DB      659 AAATGCGTAGATATGGAAGAAACACCAAGTGGGAGGAGCCTCTGGGTTAACATGTA 718
QY      723 CGCTCATGACGAAGAGTGGGAGCAACAGATTAAGTCAATCCCTGTAGTCCAGCGCT 782
DB      719 CACTGAGTGGTGGAAAGGAGTGGGAGCAACAGATTAAGTCAATCCCTGTAGTCCAGCGCT 778
QY      783 AAACGATGTAAGTATGTTGGGCTTATTAAG-GCTTGGTAACGAAGCTAACGCGGAA 841
DB      779 AAACGATGTAAGTATGTTGGGCTTATTAAGTCTTATGAGGCGACGCTAACGCAATTA 838
QY      842 GTTGACCGCTGGGAGTACGCTGCAAGATTAATACTCAAGAGATTTGACGGGAGCCG 901
DB      839 GTTGACCGCTGGGAGTACGCTGCAAGATTAATACTCAAGAGATTTGACGGGAGCCG 898
QY      902 CACAAGCGGTGATTAATGTAATTAATTCATGCAAGCGAAAGAACTTACCTTACCTTGG 961
DB      899 CACAAGCGGTGATTAATGTAATTAATTCATGCAAGCGAAAGAACTTACCTTACCTTGG 958
QY      962 ACATGTAGCAATTTTCTAGAGATAGATTAAGTG-CTTGGGAAAGCTTAACAGAGGTGCTG 1020
```

```
DB      959 ACATCCAATGAACCTTCCAGAGATGAGTGGGTGCTTCGGGAAACATTTAGAGACAGGTGCTG 1018
QY      1021 CATGCTGTCTGCAAGTCTGCTGCTGAGATGTTGGTTAAGTCCGCAACGAGCGCAAC 1080
DB      1019 CATGCTGTCTGCAAGTCTGCTGCTGAGATGTTGGTTAAGTCCGCAACGAGCGCAAC 1078
QY      1081 CTTGTCAATTAATTTGCCATCAATTT--GTTTGGGCACTTTAATGAGACTGCGGTGACAAAC 1138
DB      1079 CTTGTCTTAATTAATTTGCCATCAATTTGTTGGGCACTTTAATGAGAGACTGCGGTGACAAAC 1138
QY      1139 CCGAGGAAGTGGGAGTGAAGTCAAGTCTGATGCTGATGCTTATGAGGCTTACACAGT 1198
DB      1139 CCGAGGAAGTGGGAGTGAAGTCAAGTCTGATGCTGATGCTTATGAGGCTTACACAGT 1198
QY      1199 AATACAAATGCGCTGCAAGGTTCCCAACCCCGAGGGGAGGAGCTAATCTCAGAAAGG 1258
DB      1199 GCTACAAATGCTGCTGCAAGGTTCCCAACCCCGAGGGGAGGAGCTAATCTCAGAAAGG 1258
QY      1259 CGTGTAGTCCGGATCGAGTCTGCAACTGCACTCGGTAAGTGGATCGTATGATATC 1318
DB      1259 GATGTAGTCCGGATCGAGTCTGCAACTGCACTCGGTAAGTGGATCGTATGATATC 1318
QY      1319 GCGGATCAGCATGTCCGCGTGAATACGTTCCCGGCTTTTGAACACCGGCTGACACC 1378
DB      1319 GCGGATCAGAAATGTCCGCGTGAATACGTTCCCGGCTTTTGAACACCGGCTGACACC 1378
QY      1379 ATGGAGTGGGTTTACCAAGAACAGATAGTCTAACCCGTAAGAGGGCGCTTCCACGGT 1438
DB      1379 ATGGAGTGGGTTTACCAAGAACAGATAGTCTAACCCGTAAGAGGGCGCTTCCACGGT 1438
QY      1439 GAGATTCAATGACTGGGCTG 1457
DB      1439 GTGATTCAATGACTGGGCTG 1457

RESULT 21
US-09-821-016-5
; Sequence 5; Application US/09821016
; Patent No. 6485951
; GENERAL INFORMATION:
; APPLICANT: CANON INC.
; TITLE OF INVENTION: Polyhydroxyalkanoate Synthase and Gene Encoding the Same Enzyme
; FILE REFERENCE: 4051021
; CURRENT APPLICATION NUMBER: US/09/821,016
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii 161; BP-7376
; FEATURE:
US-09-821-016-5

Query Match      70.1%; Score 1021; DB 3; Length 1501;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 235; Indels 6; Gaps 5;

QY      3 TGAACGCTGGCGGCATGCTTTTACATGCAATGCAATGCGACAGCGAGTGCCTTGCATCT 62
DB      1 TGAACGCTGGCGGCATGCTTTTACATGCAATGCAATGCGACAGCGAGTGCCTTGCATCT 59
QY      63 GGTGGCGAGTGGGCGACGGGTGAGTATGCAATGCGAACGATTCAGAAAGAGGGGGGTAA 122
DB      60 GAAATTC-AGCGGCGAGACGGGTGAGTATGCAATGCGAACGATTCGCTGGTGGTGGGAGAAC 118
QY      123 GCATCGAAAGATGTGCTAATATCCGATATATCTTAAGSAGAAAGCAGGGAGTCGAAGA 182
DB      119 GTCTCGAAAGGAGCGCTAATATCCGATATGCTTAAGSAGAAAGCAGGGAGTCCTCGGG 178
QY      183 CCTTGGCGCTTTTGAAGCGCGCATGCTCTGATTAAGTCTGATGCTGGGTAAAGGCTTACCA 242
```

Db 179 CCTTGCCTATCAGATGAGCCTAGTGCATTAAGTTGGTGAAGTAAATGCTACCA 238  
QY 243 AGGCGACATCAGTATGTTGGTCTGAGAGACGACACGACCACTGGAGTGAACACGCG 302  
Db 239 AGGCGACATCCTTATCTGTCTGAGAGATGATCACTGACCTGGAATCGAGACGCT 298  
QY 303 CCAGACTCTTACGGGAGGACAGCACTGGGAAATTTGGACAAATGGGCGCAAGCTGATCA 362  
Db 299 CCAGACTCTTACGGGAGGACAGCACTGGGAAATTTGGACAAATGGGCGCAAGCTGATCA 358  
QY 363 GCAATGCCGCTGAGTGAAGAGCCTTCCGGTTGTAACTCTTTCAATCGAAGAA 422  
Db 359 GCCATGCCGCTGAGTGAAGAGCCTTCCGGTTGTAACTCTTTCAATCGAAGAA 418  
QY 423 AGGTACGGTAAATTAATCGTGAACCATGATCGATTCAGAGAAAGACACCGGCTAACTA 482  
Db 419 GGCATTAACTTAATTAATCGTGAACCATGATTCAGAGAAAGACACCGGCTAACTC 478  
QY 483 CGTCCAGACAGCCCGGTAATCGTAGGGTGCAGAGCGTTAATCGGAAATTAATCGGGCGTAA 542  
Db 479 TGTGCCAGACAGCCCGGTAATCAAGAGGTGCAGAGCGTTAATCGGAAATTAATCGGGCGTAA 538  
QY 543 AGGTTGGCAGAGCGGCTTGTAACTCAGATGTGAATCCCGGGCTTAACTCGGGAATG 602  
Db 539 AGCGCGGTAGGTGTGTAACTTGAATGTGAAGCCCGGGCTCAACTGGGAATCG 598  
QY 603 CGTTGAACCTACAAGAGTAGAGTGTGAGAGGAGGAGGAGGAGTTCATGATGTAGCAGTG 662  
Db 599 CATTCAAACTGACAGAGCTAGAGTATGTAGAGGAGTGTGAATTTCTGTGTAGCGGTG 658  
QY 663 AATGCGTAGAGATATGGAAGAACATCGATGGCAAGGACGCTCTGGGTTAACTACTGA 722  
Db 659 AATGCGTAGATATGGAAGAACACAGTGGCGCAAGGCGCACTGTGATCTGA 718  
QY 723 CGCTCATGACAGAAAGGTGGGAGCAAAACGATTAATCCTGTGTATGTCACGCGCT 782  
Db 719 CACTGAGGTGCAAGAGCGTGGGAGCAAAACGATTAATCCTGTGTATGTCACGCGCT 778  
QY 783 AAAGATGTCAACTAGTTGTTGGGCTTATAG-GCTTGGTAAAGAGCTAACCGGTGA 841  
Db 779 AAAGATGTCAACTAGCGGTGGAGCGCTTAACTGTAGTGGCGCACTAACCATTA 838  
QY 842 GTTGAACGCTGGGAGATACGCTGCAGATTAATACTCAAGAAATTAATGACGGGAGCCG 901  
Db 839 GTTGAACGCTGGGAGATACGCGCGCAAGTTAAACTCAATGATTAATGACGGGAGCCG 898  
QY 902 CACAAGCGGTGATATGTGATTAATTCAGATGCAACCGGAAACCTTACTTACCTTG 961  
Db 899 CACAAGCGGTGATATGTGATTAATTCAGATGCAACCGGAAACCTTACTTACCGCTTG 958  
QY 962 ACATGTACCGAATTTCTAGAGATGATTAAGT-GCTTGGGAGCGCTAACACAGGTGCTG 1020  
Db 959 ACATGTACCGAATTTCTAGAGATGATGAGTGGCTTGGGAAACTTAAGACAGGTGCTG 1018  
QY 1021 CATGAGTGTGCTCAGCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACAGCGCAAC 1080  
Db 1019 CATGAGTGTGCTCAGCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACAGCGCAAC 1078  
QY 1081 CTTGTCAATTAATGTCATCAATTT-GGTTGGGCACTTTAATGAGACTGCCGTGCAAA 1138  
Db 1079 CTTGTCAATTAATGTCATCAATTTGAGTGGGCACTTAAAGAGACTGCCGTGCAAA 1138  
QY 1139 CGGAGGAAGTGGGAGTACGTCAGATCTCAATGGCCCTTAATGGTTAGGGGCTTACAGCT 1198  
Db 1139 CGGAGGAAGTGGGAGTACGTCAGATCTCAATGGCCCTTAATGGGCTTACAGAGCT 1198  
QY 1199 AATACAATGCGCTACAGAGGTTGCAACCCGCGAGGAGGCTTAATCTCAGAAACG 1258  
Db 1199 GCTACAAATGCTCGATACAGAGGTTGCCAAGCGGAGGTGAATTCACAAAC 1258  
QY 1259 CGTGTAGTCCGATCGAGTGTGCAACTCGACTCCGTAAGTGGAAATCGTATGATTC 1318  
Db 1259 GATGTAGTCCGATCGAGTGTGCAACTCGACTCCGTAAGTGGAAATCGTATGATTC 1318

RESULT 22  
US-09-745-476-1  
; Sequence 1, Application US/09745476  
; Patent No. 6521429  
; GENERAL INFORMATION:  
; APPLICANT: CANON INC.  
; TITLE OF INVENTION: Preparation of Poly-hydroxyalkanoic Acid  
; FILE REFERENCE: 4351008  
; CURRENT APPLICATION NUMBER: US/09/745,476  
; CURRENT FILING DATE: 2000-12-26  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: Microsoft Word  
; SEQ ID NO 1  
; LENGTH: 1501  
; TYPE: DNA  
; ORGANISM: Pseudomonas jessenii Pl61 ; FERM P-17445  
US-09-745-476-1

Query Match 70.1%; Score 1021; DB 3; Length 1501;  
Best Local Similarity 83.5%; Pred. No. 0;  
Matches 1218; Conservative 0; Mismatches 235; Indels 6; Gaps 5;

QY 3 TGAACGTGGCGGCTATCTTAACATGCAAGTGCAGACGGACGACGAGTCTTGACATCT 62  
Db 1 TGAACGTGGCGGCGGCTATCTTAACATGCAAGTGCAGACGGATG-ACGGAGCTTGCTCT 59  
QY 63 GGTGGCGAGTGGCGGAGCGGTGAGTATGATCGAAGCTATCCAGAAAGGGGGTAACT 122  
Db 60 GAATTC-AGCGCGGAGCGGTGAGTATGCTAGAAATCTGCTGTATGTTGGGAGCAAC 118  
QY 123 GCATCGAAGATGTCTAATACCGCATATATCTTAAGAGAAAGCAGGGGATCGAAAGA 182  
Db 119 GTCTCGAAGAGGACGTAATACCGCATATCTCTTACGGGAAAGCAGGGGACCTTCGGG 178  
QY 183 CTTTGGCCTTTGGAGCGGCGGATGTGATTAAGTATGTTGGTGGGTTAAAGGCTTACCA 242  
Db 179 CTTTGGCCTTACAGATGAGCTTAGTGTGATTAAGTATGTTGGTGGGTTAAATGGCTTACCA 238  
QY 243 AGGCGAGATCAGATGTTGGTCTGAGAGGAGCGACACGACCACTGGGACTGAGACGCG 302  
Db 239 AGGCGAGATCAGATGTTGGTCTGAGAGATGATCACTGACCTGGAATCTGAGACAGCT 298  
QY 303 CCAGACTCTTACGGGAGGACAGCACTGGGAAATTTGGACAAATGGGCGCAAGCTGATCA 362  
Db 299 CCAGACTCTTACGGGAGGACAGCACTGGGAAATTTGGACAAATGGGCGCAAGCTGATCA 358  
QY 363 GCAATGCCGCTGAGTGAAGAGCCTTCCGGTTGTAACTCTTTCAATCGAAGAA 422  
Db 359 GCCATGCCGCTGAGTGAAGAGCCTTCCGGTTGTAACTCTTTCAATCGAAGAA 418  
QY 423 AGGTACGGTAAATTAATCGTGAACCATGATCGATTCAGAGAAAGACACCGGCTAACTA 482  
Db 419 GGCATTAACTTAATTAATCGTGAACCATGATTCAGAGAAAGACACCGGCTAACTC 478  
QY 483 CGTCCAGACAGCCCGGTAATCGTAGGGTGCAGAGCGTTAATCGGAAATTAATCGGGCGTAA 542  
Db 479 TGTGCCAGACAGCCCGGTAATCAAGAGGTGCAGAGCGTTAATCGGAAATTAATCGGGCGTAA 538  
QY 543 AGGTTGGCAGAGCGGCTTGTAACTCAGATGTGAATCCCGGGCTTAACTCGGGAATG 602

```

Db      539 AGCGCGGTAGGTGTTTGTAACTGTAAGTGAAGAACCCCGGCTCAACCTGGGAACCTG 598
Qy      603 CGTTTGAACCTACAAACCTAGAGTGGCAAGGAGGTGAATTCATGTGTAGCAAGT 662
Db      599 CATTCAAACTGACAACTAGAGTGTAGAGGGGTGTGAATTTCTGTGTAGCGGTG 658
Qy      663 AAATGCTAGAGATATGGAAGAACATGATGGCCAAAGCCCTCTGGGTAAACATGA 722
Db      659 AAATGCTAGATATGGAAGAACACAGTGGCCAAAGCCCAACCTGTGACTGATCTGA 718
Qy      723 CGCTCATGCAAGAAAGGTGGGAGCAAAACAGATTAGATACCTGTGTAGTCCAGCCCT 782
Db      719 CACTGAGGTGCGAAAGGTGGGAGCAAAACAGATTAGATACCTGTGTAGTCCAGCCCT 778
Qy      783 AAACGATGTCAACTAGTGTGTGGCCCTTATTAG-GCTTGTGAACGAAGCTAACGCGTAA 841
Db      779 AAACGATGTCAACTAGCCGTTGGGAGCCTTGAAGCTTATGTGGCGAGCTAACGATTA 838
Qy      842 GTTGAACCGCTGGGAGTACGCTGCGAAAGATTAAACTCAAGGAATTGACGGGGAACCG 901
Db      839 GTTGAACCGCTGGGAGTACGCGCAAGGTAAACTCAAGTGAACGGGGGCCG 898
Qy      902 CACAAGCGGTGATTAATGATTAATTCATGCAAGCGCAAAACCTTACTTACCCCTTG 961
Db      899 CACAAGCGGTGAGCATGTGTAAATTCAGAGCAACCGAAGAACTTACAGGCTTG 958
Qy      962 ACATGTACGAAATTTTGTAGAGATAGATTAGTG-CTTGGGAACGCTAACAGGTCTG 1020
Db      959 ACATGCAATGAATTTTCAAGATGATGGGTGCTTGGGAACATTTGAGACAGGTCTG 1018
Qy      1021 CATGGCTGTGTCAGCTGCTGTGTGAGATGTTGGGTTAAGTCCCGAAGAGCGCAAC 1080
Db      1019 CATGGCTGTGTCAGCTGCTGTGTGAGATGTTGGGTTAAGTCCCGAAGAGCGCAAC 1078
Qy      1081 CTGTGCTAATATTGCGCATCTTT--GGTGGGCACTTTAATGAGACTGCGGTGACAAAC 1138
Db      1079 CTGTGCTTATGTTACAGCACTGATATGTGGGCACTTTAAGAGACTGCGGTGACAAAC 1138
Qy      1139 CGAGGAAGGTGGGATGAGCTCAAGTCTCATGGCCCTTATGGGTGAGGCTTACACGT 1198
Db      1139 CGAGGAAGGTGGGATGAGCTCAAGTCTCATGGCCCTTATGGGTGAGGCTTACACGT 1198
Qy      1139 AATTAATGCGGCTACAGAGGTTGCCAACCCGCGAGGGGAGCTATCTCAAGAACG 1258
Db      1199 GCTAACATGTGCTGCTACAGAGGTTGCCAACCCGCGAGGTGAGCTAATCCCAAAAC 1258
Qy      1259 CGTGTGATCGGATCGGAGTCTGCACTCGTGAAGTCCGTAATCGTAATGCTATG 1318
Db      1259 GATGTGATCGGATCGGAGTCTGCACTCGTGAAGTCCGTAATCGTAATGCTATG 1318
Qy      1319 GCGGATCAGCATGTGCGGTGAAATACGTTCCGGGTCTTGTACACACGCGCGCTACACC 1378
Db      1319 GCGGATCAGCATGTGCGGTGAAATACGTTCCGGGTCTTGTACACACGCGCGCTACACC 1378
Qy      1379 ATGGGAGTGGGTTTACCAAGAGAGTGTCTTACCGTTAAGAGAGGCGCTTGCACGGT 1438
Db      1379 ATGGGAGTGGGTTTACCAAGAGAGTGTCTTACCGTTAAGAGAGGCGCTTGCACGGT 1438
Qy      1439 GAGATTCATGACGCGGGTG 1457
Db      1439 GTGATTCATGACGCGGGTG 1457
```

## RESULT 23

```
US-09-748-205-1
; Sequence 1, Application US/09748205
```

```
; Patent No. 6586562
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Canon Inc.
```

```
; TITLE OF INVENTION: Poly(hydroxyalkanoate)ib manufacturing method, and microorganism
```

```
; TITLE OF INVENTION: those are used for the method.
```

```
; FILE REFERENCE: 4351009
```

```

; CURRENT APPLICATION NUMBER: US/09/748,205
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii 161 strain.
US-09-748-205-1

Query Match      70.1%; Score 1021; DB 3; Length 1501;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 235; Indels 6; Gaps 5;

Qy      3 TGAACGCTGGCGGCACTTTTACATGCAAGTCGAAGCGGACGAGTGTGCATCT 62
Db      1 TGAACGCTGGCGGCACTTTTACATGCAAGTCGAAGTCGAAGTCGAGTGTGCATCT 59
Qy      63 GGTGCGAGTGGCGGAGCGGTGATGATCATCGAACCTATCCAGAAAGGGGGTAAAC 122
Db      60 GAAATTC-AACGGCGGACGGGTGATGATGCTAGGAATCTGCTGTGTAGTGGGGGCAAC 118
Qy      123 GCATGAAAGATGTCTAATACCGCATATCTCTAAGAGAAAGCAGGGGATCGAAAGA 182
Db      119 GTCTGAAAGGAGCGCTAATACCGCATATCGTCTACGGGAGAAAGCAGGGGACCTTCGGG 178
Qy      183 CCTTGGCTTTTGGAGCGGCGCATGTCGATTAAGTATGTTGGGTAAAGGCGTACCA 242
Db      179 CCTTGGCTTTTGGAGCGGCGCATGTCGATTAAGTATGTTGGGTAAAGGCGTACCA 238
Qy      243 AGCGCAGATCAGTATGTTGTTGCTGAGAGGACGACGACCACTGGGACTGAGACACGCG 302
Db      239 AGCGCAGATCAGTATGTTGTTGCTGAGAGGATGATCAGTACACTGGAACCTGAGACCGGT 298
Qy      303 CCAAGCTCTTACGGGAGGACGACGATGGGGAATTTTGGACAATGGGCGCAAGCTGATCCA 362
Db      299 CCAAGCTCTTACGGGAGGACGACGATGGGGAATTTTGGACAATGGGCGCAAGCTGATCCA 358
Qy      363 GCAATGCCCGGTGAGTGAAGGAAGGCTTCGGGTGTAAGCTCTTTCAGTGGAGAAAG 422
Db      359 GCAATGCCCGGTGAGTGAAGGAAGGCTTCGGGTGTAAGCTCTTTCAGTGGAGAAAG 418
Qy      423 AGGTTCGCTAATTAATCGTGACCATGACGCTATCGACAGAAAGACCCGCTTAACCTA 482
Db      419 AGGTTCGCTAATTAATCGTGACCATGACGCTATCGACAGAAAGACCCGCTTAACCTA 478
Qy      483 CGTGCAGAGCCGCGGTAAATAGTGGGTGCAAGGCTTAATCGGAATTACTGGGCGTAA 542
Db      479 TGTGCCAGAGCCGCGGTAAATAGTGGGTGCAAGGCTTAATCGGAATTACTGGGCGTAA 538
Qy      543 AGGTGCGGAGCGCGCTTGTAAATGTCAGATGTGAAATCCCGGGGCTTAACCTGGGAATTG 602
Db      539 AGGTGCGGAGCGCGCTTGTAAATGTCAGATGTGAAATCCCGGGGCTTAACCTGGGAATTG 598
Qy      603 CGTTTGAACCTACAAAGCTAGAGTGGCAGAGGAGGTGAATTCATGTGTAGCAAGT 662
Db      599 CATTCAAACTGACAACTAGAGTGTAGAGGGGTGTGAATTTCTGTGTAGCGGTG 658
Qy      663 AAATGCTAGAGATATGGAAGAACATGATGGCCAAAGCCCTCTGGGTAAACATGA 722
Db      659 AAATGCTAGATATGGAAGAACACAGTGGCCAAAGCCCAACCTGTGACTGATCTGA 718
Qy      723 CGCTCATGCAAGAAAGGTGGGAGCAAAACAGATTAGATACCTGTGTAGTCCAGCCCT 782
Db      719 CACTGAGGTGCGAAAGGTGGGAGCAAAACAGATTAGATACCTGTGTAGTCCAGCCCT 778
Qy      783 AAACGATGTCAACTAGTGTGTGGCCCTTATTAG-GCTTGTGAACGAAGCTAACGCGTAA 841
Db      779 AAACGATGTCAACTAGCCGTTGGGAGCCTTGAAGCTTATGTGGCGAGCTAACGATTA 838
Qy      842 GTTGAACCGCTGGGAGTACGCTGCGAAAGATTAAACTCAAGGAATTGACGGGGAACCG 901
Db      839 GTTGAACCGCTGGGAGTACGCGCAAGGTAAACTCAAGTGAATGAGCGGGGCCG 898
```

Qy		902	CACAGCGGCTGGATTAATGTGATTAATTCCATGCAAGCGAAAAAACCCTTACCTC	961
Db		899	CACAAAGGCTGGAGCATGTGGTTAAATTGCAGAACGCAAGAACCCTTACAGGCCCTTG	958
Qy		962	ACAATGTAGCGAATTTTCTTAGAGATAGATTAGTG-CTTCGGGAACGCTTAACAGAGTCTG	1020
Db		959	ACATTCATAATACCTTTCCAGAGATAGATGGGTGCTTCGGGAACATTGAGACAGGTCTG	1018
Qy		1021	CATGGCTGTGCTCAGCTCGTGTCTGTAGATGTTGGTTAAGTCCCGCAAGAGCGCAAC	1080
Db		1019	CATGGCTGTGCTCAGCTCGTGTCTGTAGATGTTGGTTAAGTCCCGTAACGAGCGCAAC	1078
Qy		1081	CTTGTCAATTAATTCCTCATATT--GGTGGGCACTTTAANTGACTCGSGTGACAAC	1138
Db		1079	CTTGTCTTATGTTAACAGACAGTAATGCTGGGCACTCTAAGAGACTGCGGTGACAAC	1138
Qy		1139	CGAGGAAGGCTGGGATGACGTCAAGTCCCTCATGGCCCTTAATGGGTAGGGCTTACACGT	1198
Db		1139	CGAGGAAGGCTGGGATGACGTCAAGTCAATCATATGCCCCCTTACGCCCTGGGTACACAGT	1198
Qy		1199	AATCAAATGGCGCGCTACAGAGGCTTGCACAACCGCGAGGGGAGCTAATCTCAGAAAACG	1258
Db		1199	GCTACAAATGGCTGGTACAGAGGCTTGCACAACCGCGAGGTGAGACTAATCCCAAAAAC	1258
Qy		1259	CGTGTGATCGCGGATCGGATCTGCACTCGACTCCGTGAAGTGGGAATGCGTATGTAATC	1318
Db		1259	GATGTGATCGCGGATCTGCACTCGACTCGGGAATGCGGAATGCTATGTAATC	1318
Qy		1319	GCGGATCAGACATGCGGGGTGAATACGTTCCGGGCTCTTGTACACACCGCGCTCACAC	1378
Db		1319	GCGAATCAGAAATGTCGGGTGAATACGTTCCCGGGCTTGTACACACCGCGCTCACAC	1378
Qy		1379	ATGGAGTGGGTTTCAACCAAGACAGGTAGTCTAAACCGTAAGAGGGCGCTTGCACGCT	1438
Db		1379	ATGGAGTGGGTTTCAACCAAGATAGTACTAATCTTCGGAGAGACGGTTACACGCGT	1438
Qy		1439	GAGATTCACTACTGGGGTG	1457
Db		1439	GTGATTCACTACTGGGGTG	1457
RESULT 24				
US-09-951-720-1				
; Sequence 1, Application US/09951720				
; Patent No. 6635782				
; GENERAL INFORMATION:				
; APPLICANT: Canon Kabushiki Kaisha				
; TITLE OF INVENTION: Polymyxinylkanonate and Manufacturing Method Thereof				
; FILE REFERENCE: 4477001				
; CURRENT APPLICATION NUMBER: US/09/951.720				
; PRIORITY FILING DATE: 2000-09-14				
; PRIOR APPLICATION NUMBER: JP 279900/2000				
; JP 378827/2000				
; JP 165238/2001				
; JP 165509/2001				
; JP 275063/2001				
; PRIORITY FILING DATE: 2000-09-14				
; 2000-12-13				
; 2001-05-31				
; 2001-05-31				
; 2001-09-11				
; NUMBER OF SEQ ID NOS: 1				
; SEQ ID NO 1				
; LENGTH: 1501				
; TYPE: DNA				
; ORGANISM: Pseudomonas jessenii P161 strain.				
US-09-951-720-1				
Qy	Match	70.1%;	Score 1021;	DB 3; Length 1501;
	Best Local Similarity	83.5%;	Pred. No. 0;	
	Matches 118;	Conservative	0; Mismatches 235;	Indels 6; Gaps 5;
Qy	3 TGAACGCTGGCGGCAATGCTTTACACATGCAAGTCGACGCGACGACGATGCTTGACATCT 62			

[illegible]

Db 1079 CTTGTCTTACGACGACGTAATGTGGGCACTTAAGAGACTGCCGTGACAAAC 1138  
QY 1139 CGAGGAAAGGTGGGATGACGTCAAGTCTTCATGSCCTTATGGGTAAGGCTTCAACGT 1198  
Db 1139 CGAGGAAAGGTGGGATGACGTCAAGTCAATGACCTTACGCGCTGAGCTACACAGT 1198  
QY 1199 AATACATGCGGGTACAGAGGGTGGCCAAACCGCGAGGGGAGCTAATCTGAAAGG 1258  
Db 1199 GCTACATGCTCGGTACAGAGGGTGGCCAAACCGCGAGGGTGAAGCTAATCCCAAAAC 1258  
QY 1259 CGTGTAGTCCGATCGGATCTGCAACTCGACTCCGTGAAGTCGGAATCGTAATATC 1318  
Db 1259 GATGTAGTCCGATCGGATCTGCAACTCGACTCCGTGAAGTCGGAATCGTAATATC 1318  
QY 1319 GCGGATCAGCATGTCCGGTGAATACGTTCCGGGTCTTGTACACACCGCCGTCAACC 1378  
Db 1319 GCGAATCAGATGTCCGGTGAATACGTTCCGGGTCTTGTACACACCGCCGTCAACC 1378  
QY 1379 ATGGAGTGGGTTTACACAGAGAGTACTTACCGCTAAGAGGGCGCTTCCACCGT 1438  
Db 1379 ATGGAGTGGGTTTACACAGAGAGTACTTACCGCTAAGAGGGCGCTTCCACCGT 1438  
QY 1439 GAGATTCATGACTGGGGTG 1457  
Db 1439 GTGATTCATGACTGGGGTG 1457

RESULT 25  
US-10-411-319-1  
; Sequence 1, Application US/10411319  
; Patent No. 664381  
; GENERAL INFORMATION:  
; APPLICANT: Canon Inc.  
; TITLE OF INVENTION: Polyhydroxynate, Method For Production Thereof And Microorganisms  
; FILE REFERENCE: 03500.015001.1  
; CURRENT APPLICATION NUMBER: US/10/411.319  
; PRIORITY FILING DATE: 2003-04-11  
; PRIOR APPLICATION NUMBER: US 09/748,205  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 1  
; LENGTH: 1501  
; TYPE: DNA  
; ORGANISM: Pseudomonas jessenii 161 strain  
US-10-411-319-1

Query Match 70.1%; Score 1021; DB 3; Length 1501;  
Best Local Similarity 83.5%; Pred. No. 0;  
Matches 1218; Conservative 0; Mismatches 235; Indels 6; Gaps 5;

QY 3 TGAACGCTGCGGCGATCTTACCATGCAATGCAAGCGGACGACGATGCTTGACATC 62  
Db 1 TGAACGCTGCGGCGGCGCTTACCATGCAATGCAAGCGGATG-ACGGAGCTTGCTCT 59  
QY 63 GGTGGCGAGTGGCGGAGCGGGTGAATATGCATCGGAACGTATCAGAAAGAGGGGGTAA 122  
Db 60 GAATTC-AGCGGCGGAGCGGGTGAATATGCCTAGGAATCTGCTGGTATGGGGGACAAC 118  
QY 123 GCATTCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATCGAAAG 182  
Db 119 GTCCTGAAAGGAGCGCTAATACCGCATATCTTACCGGAGAAAGCAGGGGACCTTCGG 178  
QY 183 CTTTGCGCTTTGGAGGCGCGATGCTGATTTAGTCTGATTTGTTGGGGTAAAGGCTTACCA 242  
Db 179 CTTTGCGCTTACGATGATGCTTATAGTCTGATTTAGTCTGATTTGTTGGGGTAAAGGCTTACCA 228  
QY 243 AGCGAGCATCACTAGTGTGCTGAGAGAGCAACGACCACTGGAGCTGAGACACGCG 302  
Db 239 AGCGAGCATCCGTAATCTGTGTGAGAGAGTGAATCACTGGAATGAGACACGCGT 298  
QY 303 CCAAGATCTCTACGAGAGGACAGCTGGGGAATTTTGAACAATGGGCGCAAGCCTGATCCA 362

Db 299 CCAAGATCTCTACGAGAGGACAGCTGGGGAATTTTGAACAATGGGCGAAAGCCTGATCCA 358  
QY 363 GCAATGCCGCTGATGAAAGAGGCTTGGGTTGAAAGCTTTCACTGAGAGAGAA 422  
Db 359 GCCATGCCGCTGATGAAAGAGGCTTGGGTTGAAAGGCTTTAAGGCTTTAAGTTGGAGAG 418  
QY 423 AGGTTACGTTAATATCGTGACCCATGACGCTATCGACAGAGAGAGCAGGCTTACTA 482  
Db 419 GGCATTAACCTTAATACGTTAGTTGACGTTACCGAGATAGAACCGGCTTACTC 478  
QY 483 CGTGCCAGAGCGCGGTAATACGTAGGTTGCAACGTTAATCGAATTTACTGGCGTAA 542  
Db 479 TGTGCCAGAGCGCGGTAATACAGAGGTTGCAACGTTAATCGAATTTACTGGCGTAA 538  
QY 543 AGGGTGGCAGGCGGCTTTGAAAGTCAAGTGAATCCCGGGCTTAACTTGGAAATG 602  
Db 539 AGCGCGGTAAGGTTGTTGTTGATGATGAAAGCCCGGGCTCAACCTGGGAACTG 598  
QY 603 CGTTGAAACTCAAAAGCTAAGTGTGGCAGAGGAGTGAATTCATGTGTACAGATG 662  
Db 599 CATTCAAACTGACAAAGCTAAGATGTGAGAGGGTGTGGAATTTCTGTGTACGCGG 658  
QY 663 AATGCGTAGATATGGAAGACATCGATGCGAAGGCGCTCGGAGTTAACACTGA 722  
Db 659 AATGCGTAGATATGGAAGAGAACACAGATGCGAAGGCGACCACTGAGTATCTGA 718  
QY 723 CGCTACATGACAAAGCGTGGGAGCAACAGATTAATACCTGTTATCCAGCCCT 782  
Db 719 CACTAGGTGCGAAAGCGTGGGAGCAACAGATTAATACCTGTTATGTCACGCGCT 778  
QY 783 AATGATGCTCACTAGTTTGGGCTTATG-CTTGTGTAACGAGCTAACGGGTAA 841  
Db 779 AAACGATGCACTAGCTTGGGCTTATG-CTTGTGTAACGAGCTAACGGGTAA 838  
QY 842 GTTGACCGGCTGGGAGTACGCTCGCAAGATTAATACTCAAAAGATTGACGGGACCG 901  
Db 839 GTTGACCGGCTGGGAGTACGCTCGCAAGATTAATACTCAAAAGATTGACGGGACCG 898  
QY 902 CACAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAAACCTTACCTACCTTG 961  
Db 899 CACAAGCGGTGAGATGTGTTAATTCGAAGCAACGCGAAAACCTTACCTACCTTG 958  
QY 962 ACATGAGGAAATTTTCTAGATATGATAGTG-CTTGGGAAACCTTAACAAGTGCTG 1020  
Db 959 ACATGAGGAAATTTTCTAGATATGATAGTG-CTTGGGAAACCTTAACAAGTGCTG 1018  
QY 1021 CATGCTGCTGACAGTCTGTCGTGAGATGTTGGGTTAAGTCCGCAACGACGCAAC 1080  
Db 1019 CATGCTGCTGACAGTCTGTCGTGAGATGTTGGGTTAAGTCCGTAACGACGCAAC 1078  
QY 1081 CTTGTCTTAATTTGCCATCAATTT--GGTGGCACTTTAATGAGCTGCGGTGACAAAC 1138  
Db 1079 CTTGTCTTAATTTGCCATCAATTT--GGTGGCACTTTAATGAGAGCTGCGGTGACAAAC 1138  
QY 1139 CGAGGAAAGTGGGATGACGTCAAGTCTCATGCGCTTATGGTATGGGCTTACACAGT 1198  
Db 1139 CGAGGAAAGTGGGATGACGTCAAGTCTCATGCGCTTATGGGCTTACACACAGT 1198  
QY 1199 AATACATGCGCGTACAGAGGGTGGCCAAACCGCGAGGGGAGCTAATCTCAGAAAGG 1258  
Db 1199 GCTACATGCTCGGTACAGAGGGTGGCCAAACCGCGAGGGTGAAGCTAATCCCAAAAC 1258  
QY 1259 CGTGTAGTCCGATCGGATCTGCAACTCGACTCCGTGAAGTCGGAATCGTATGTAATC 1318  
Db 1259 GATGTAGTCCGATCGGATCTGCAACTCGACTCCGTGAAGTCGGAATCGTATGTAATC 1318  
QY 1319 GCGAATCAGCATGTCCGGTGAATACGTTCCGGGTCTTGTACACACCGCCGTCAACC 1378  
Db 1319 GCGAATCAGAAATGTCCGGTGAATACGTTCCGGGTCTTGTACACACCGCCGTCAACC 1378  
QY 1379 ATGGAGTGGGTTTACCAAGAGAGGATAGTAACTCGTAACGTAAGAGGGCGCTTGACAGGT 1438





```
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: JP 2000-095004
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii P161 ; BP-7376
; FEATURE:
; FEATURE: cDNA to 16S rRNA
US-10-266-787-5

Query Match      70.1%; Score 1021; DB 3; Length 1501;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 235; Indels 6; Gaps 5;

QY      3 TGAACGCTGGCGGCATCTTTACACATGCAAGTCGAACGCGACACCGATGCTTGCATCT 62
DB      1 TGAACGCTGGCGGCATCTTTACACATGCAAGTCGAACGCGATG-ACGCGAGCTTGCCTCT 59

QY      63 GGTGGCGAGTGGCGGACGCGGTGAATATGATCGGAACGTATCCAGAAAGGGGGGTAC 122
DB      60 GAATTC-AGCGCGGACGCGGTGAATATGCTAGGAATCTGCTGTTAGTGGGGGACAAAC 118

QY      123 GCATCGAAAGATGCTTAATACCGCATATCTTAAGAGAGAAAGACGAGGATCGAAAGA 182
DB      119 GTCTCGAAAGGACGCTTAATACCGCATATCTCTACCGAGAAAGACGGGACCTTCGGG 178

QY      183 CCTTGGCTTTTGAAGCGCGCATATGCTGATTTAGTCTGATTTGTTGGGTAAAGGCTTACCA 242
DB      179 CCTTGGCTTTTGAAGCGCGCATATGCTGATTTAGTCTGATTTGTTGGGTAAAGGCTTACCA 238

QY      243 AGCGCAGATCACTAGTGTGTCGAGAGACGACCGACACTGGAGCTGAGACACGCGC 302
DB      239 AGCGCAGATCCGTAATCTGATGTCGAGAGATGATCACTGACACTGGAATGAGACACGCGT 298

QY      303 CCAGATCTCTTACCGGAGGACGACGAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATCCA 362
DB      299 CCAGATCTCTTACCGGAGGACGAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATCCA 358

QY      363 GCAATGCCGCGGTAGTGAAGAAAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGAGAA 422
DB      359 GCCATGCCGCGGTAGTGAAGAAAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGAGAAAG 418

QY      423 AGGTACCGTAAATTAATCGTGAACCATGACGGTATGACAGAAAGACCGGCTAACTA 482
DB      419 GGCATTAACTTAATTAATCGTGAAGCATTAACGACAGAAATTAAGCACCGGCTAACTC 478

QY      483 CGTCCAGACGCGCGGCTAATACGTAGGGTGCAGAGGCTTAATCGGAATTACTGGGCGTAA 542
DB      479 TGTCCAGACGCGCGGCTAATACAGAGGGTGCAGAGGCTTAATCGGAATTACTGGGCGTAA 538

QY      543 AGGGTGGCAGCGCGGCTTGAATGATGATGTAATCCCGGGCTTAACTTGGGAATTTG 602
DB      539 AGGGTGGCAGCGCGGCTTGAATGATGATGTAATCCCGGGCTTAACTTGGGAATTTG 598

QY      603 CGTTGAAACTACAAAGCTAGTGTGCGACAGAGGAGTGAATTCATGTTGAGAGTG 662
DB      599 CATTCAAAACTGACAAAGCTAGTGTGCGACAGAGGAGTGAATTCATGTTGAGAGTG 658

QY      663 AAATGCGTAGAGATATGGAAGAACTGATGCGGAAGGACCTCTCTGGTTAACTGTA 722
DB      659 AAATGCGTAGATATGGAAGAAACACCAAGTGGGAAAGGACCACTTGAATATCTGTA 718

QY      723 CGCTCAAGCAGAAAGGTGGGAGCAACAGGATTAAGATACCTGTTAGTCCAGCGCT 782
DB      719 CACTGAGGTGCGAAAGGTGGGAGCAACAGGATTAAGATACCTGTTAGTCCAGCGCT 778

QY      783 AAACGATGCACTAGTGTGGCTTATTAAG-GCTTGGTAAAGAACTGAACGCGTGA 841
DB      779 AAACGATGCACTAGCGCTTGGGAGCTTGAAGCTTGAATGAGCGACGTAAGCAATTA 838
```

```
QY      842 GTTGAACCGGCTGGGAGTACGCGTCCGAGATTTAAACTCAAAAGATTTAGACGGGACCG 901
DB      839 GTTGAACCGGCTGGGAGTACGCGTCCGAGATTTAAACTCAAAAGATTTAGACGGGACCG 898

QY      902 CACAACCGGTGATTAATGATTAATTCGATGCAACGCGAAACCTTACCTACCTTTG 961
DB      899 CACAACCGGTGATGATGTTTAAATTCGAAAGCAACGGAAGAACTTACCAAGGCTTTG 958

QY      962 ACATGATGGAATTTTCTTAAGATATGATGAG-CTTGGGGAACCTTAACAGATGCTG 1020
DB      959 ACATGATGGAATTTTCTTAAGATATGATGAG-CTTGGGGAACCTTAAGAGATGCTG 1018

QY      1021 CATGCTGCTGACGCTGCTGATGATGATGTTGGGTAAAGTCCCGCAACGAGCCCAAC 1080
DB      1019 CATGCTGCTGACGCTGCTGATGATGATGTTGGGTAAAGTCCCGCAACGAGCCCAAC 1078

QY      1081 CTGTGATTAATTTGCCATCAATTT--GGTTGGGACCTTTAATGAGACTGCGGATGACAAAC 1138
DB      1079 CTGTGCTTAATTTACAGACGTAATGATGAGGACCTTAAGAGACTGCGGATGACAAAC 1138

QY      1139 CCGAGGAAGTGGGATGACGTCAGTCTCATTTGACCTTAAGGTTAGGGCTTCAACGCT 1198
DB      1139 CCGAGGAAGTGGGATGACGTCAGTCTCATTTGACCTTAAGGCTTCAACGAGCTTCAACGCT 1198

QY      1199 AATACAAATGCGCGTACAGAGGGTTGCCAACCGCGAGGGGAGGCTTAATCTCAGAAAGCG 1258
DB      1199 GCTACAAATGCTGCTACAGAGGGTTGCCAACCGCGAGGGGAGGCTTAATCTCAGAAAGCG 1258

QY      1259 CGTGTAGTCCGAGATCGAGTCTGCAACTCGACTCGTGAAGTGGAAATCGCTAGTAATC 1318
DB      1259 GATCGTAGTCCGAGATCGAGTCTGCAACTCGACTCGTGAAGTGGAAATCGCTAGTAATC 1318

QY      1319 GCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGGCTTTTGAACACCGCCCGTACACC 1378
DB      1319 GCGGATCAGGAATGTGCGCGGTGAATACGTTCCCGGGCTTTTGAACACCGCCCGTACACC 1378

QY      1379 ATGGAGTGGGTTTACACGAAAGCAGTAGTCTAACCGTAAGAGAGGGGCTTGCACGAT 1438
DB      1379 ATGGAGTGGGTTTACACGAAAGTACGTAAGTCTTAACCTTGGAGAGAGCGTTACACGAT 1438

QY      1439 GAGATTCATGACTGGGGTG 1457
DB      1439 GTGATTCATGACTGGGGTG 1457

RESULT 28
US-09-791-610-1
; Sequence 1, Application US/09791610
; Patent No. 6861550
; GENERAL INFORMATION:
; APPLICANT: Canon Inc.
; TITLE OF INVENTION: Polylhydroxylalkanoate containing 3-hydroxybenzoylalkanoic acid as
; FILE REFERENCE: monomer unit, and method for producing the same.
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii 161 strain.
US-09-791-610-1

Query Match      70.1%; Score 1021; DB 3; Length 1501;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 235; Indels 6; Gaps 5;

QY      3 TGAACGCTGGCGGCATCTTTACACATGCAAGTCGAACGCGACACCGATGCTTGCATCT 62
DB      1 TGAACGCTGGCGGCATCTTTACACATGCAAGTCGAACGCGATG-ACGCGAGCTTGCCTCT 59

QY      63 GGTGGCGAGTGGCGGACGCGGTGAATATGATCGGAACGTATCCAGAAAGGGGGGTAC 122
```



```

Db      60 GAATTC-AGCGGCGGAGCGGTGAAGTAATGCTAGAAATCTGCTGGTAGTGGGGGACAC 118
Qy      123 GCATCGAAGATGTTGTTATACCGCATATATCTTAAGAGAGAAAGACAGGGATCGAAGA 182
Db      119 GTCTCGAAGAGGACGCTAATACCGCATACGCTCTACGGAGAGAAAGACAGGGGACCTTGGG 178
Qy      183 CTTGCGCTTTTGGAGCGGCGGATGTCTGATTAGCTAGTTAGTGGGGTAAAGGCTTACCA 242
Db      179 CTTGGGCTATCAGATAGAGCTAGGCTAGGATAGCTAGTTAGTGGTAAATGGCTTACCA 238
Qy      243 AGGCGACATCATAGTATGTTGGTCTGAGAGAGACACACCACTGGGACCTGAGACAGGC 302
Db      239 AGGCGAGATCCGTAACCTGCTGAGAGAGATCACTCACTGGAACCTGAGACAGCT 298
Qy      303 CCAAGACTCTACGGGAGGACAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATCCA 362
Db      299 CCAAGACTCTACGGGAGGACAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATCCA 358
Qy      363 GCAATGCGCGGTAGTGAAGAGGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAGAAA 422
Db      359 GCCATGCGCGGTGTGTGAAGAGGCTTCGGATTGTAAAGCACTTAAAGTTGGAGGAG 418
Qy      423 AGTTACGTAATTAATCGTGAACCAATGACGCTATCCAGAGAGAAAGACCGGCTAATCA 482
Db      419 GGCATTAACCTAATACGTTAGTGTATTGACGTTACCGACAGAAATAGACCGGCTAATC 478
Qy      483 CGTCCAGACACCGCGGTAATACGTAAGGAGGTCAGAGCGTTAATCGGAATTACTGGGCGTAA 542
Db      479 TGTCCAGACACCGCGGTAATACGAGAGGTCAGAGCGTTAATCGGAATTACTGGGCGTAA 538
Qy      543 AGGATGCGCAGCGCGCTTGTAGTCAAGTGTGAATCCCGGGCTTAACCTGGGAATTG 602
Db      539 AGCGCGCTAGTGGTTGTATAGTTGATGTGAAGAGCCCGGGCTCAACCTGGGAATTG 598
Qy      603 CGTTGAACCTACAAAGCTAGAGTGTGCAAGAGGAGTGAATTCATGTGTAGCAATG 662
Db      599 CATTCAAACTGACAAAGCTAGAGTGTGAGAGGCTGTGTGAATTTCTGTGTAGCGGTG 658
Qy      663 AAATGCGTAGAGATATGGAAGCAATCGATGGCGAAGGACGCTCTGGGTTAACTGACTGA 722
Db      659 AAATGCGTAGAGATATGGAAGCAATCGATGGCGAAGGACCACTTGAAGTGAATCTGACTGA 718
Qy      723 CGCTCATGCAAGAAAGCTGTGGGAGCAAAACAGATTAGATACCTGTAGTCCAGCCCT 782
Db      719 CACTGAGGTGCGAAGAGCTGTGGGAGCAAAACAGATTAGATACCTGTAGTCCAGCCCT 778
Qy      783 AAACGATGTCAAATGATTGTTGGGCTTAATTAG-GCTTGTATACGAACTTAAACGCTGAA 841
Db      779 AAACGATGTCAAATGATTGTTGGGAGCTTGTAGCTCTTAAGTGGCCAGCTTAAACGCTTAA 838
Qy      842 GTTGAACCGCTGTGGGAGATACGCTCCCAAGATTAAACCTCAAGAAATTGACGGGGACCG 901
Db      839 GTTGAACCGCTGTGGGAGATACGCTCCCAAGATTAAACCTCAAGAAATTGACGGGGACCG 898
Qy      902 CACAAGCGGATGGAATTAATGATTAATTCAGATGCAACGGAAGAAACCTTACTCACTTGTG 961
Db      899 CACAAGCGGATGGAATGATGTTTAATTCAGAGCAACGGAAGAAACCTTACTCACTTGTG 958
Qy      962 ACATGTAGCGAATTTTCTAGAGATAGATTAGTG-CTTGGGGAACGCTTAAACAGGTGCTG 1020
Db      959 ACATGTAGCGAATTTTCTAGAGATAGATTAGTGCTTCCGGAACATTTGAGAGAGGTGCTG 1018
Qy      1021 CATGCTGTGCTCACTGCTGTGCTGAGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1080
Db      1019 CATGCTGTGCTCACTGCTGTGCTGAGATGTTGGGTTAAGTCCGTTAAGAGCGCAAC 1078
Qy      1081 CTTGTCAATTAATGTCATCAATTT--GTTGGGCACTTTAATGAGACTGCCGATGCAAAAC 1138
Db      1079 CTTGTCTCTTAATGTTACAGACATTAATGTTGGGCACTTAAAGAGACTGCCGATGCAAAAC 1138
Qy      1139 CGAGAGAGGTGGGAGTGAAGTCTCATGTCCTTATGGCCCTTAATGGGATGGGCTTCAACGCT 1198
Db      1139 CGAGAGAGGTGGGAGTGAAGTCTCATGTCCTTATGGCCCTTAAAGGCTGGGCTTAAACAGCT 1198

```

```

Qy      1199 AATACATGCGCGCTACAGAGGTTGCCAACCCGCGAGAGGGGAGCTAATCTCAGAAAGCG 1258
Db      1199 GCTTACATGCTCGGTACAGAGGTTGCCAAGCCGCGAGGTGGAGCTAATCTCAGAAAC 1258
Qy      1259 CGTGTAGTCCGATCGGATCGAGTGTGCAACTGCACTCCGTGAGTGGAAATCGTATGATC 1318
Db      1259 GATGTAGTCCGATCGGATCGAGTGTGCAACTGCACTCCGTGAGTGGAAATCGTATGATC 1318
Qy      1319 GCGATACAGCATGTCGCGGTGAATACGTTCCGGGTTCTGTACACACCGCCGTACAC 1378
Db      1319 GCGAATCAGAAATGTCGCGGTGAATACGTTCCGGGTTCTGTACACACCGCCGTACAC 1378
Qy      1379 ATGGAGTGGGTTTACACAGACAGATGATCTTAACGTTAAGAGAGGCGCTTGCACAGT 1438
Db      1379 ATGGAGTGGGTTTACACAGACAGATGATCTTAACGTTAAGAGAGGCGCTTGCACAGT 1438
Qy      1439 GAGATTGATGACTGGGGTG 1457
Db      1439 GTGATTCATGACTGGGGTG 1457

RESULT 29
US-10-252-518-5
; Sequence 5, Application US/10252518
; Patent No. 6875596
; GENERAL INFORMATION:
; APPLICANT: Yano, Tetsuya
; APPLICANT: Imamura, Takeshi
; APPLICANT: Suda, Sakae
; APPLICANT: Horima, Tsutomu
; TITLE OF INVENTION: Polynhydcoxalkanoate Synthase and Gene Encoding the Same Enzyme
; FILE REFERENCE: 03500.015225.2
; CURRENT APPLICATION NUMBER: US/10/252,518
; PRIORITY FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: JP 2000-095004
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii Pl61 ; Bp-7376
; FEATURE:
; ; FEATURE: cDNA to 16S rRNA
US-10-252-518-5

Query Match      70.1%; Score 1021; DB 3; Length 1501;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 235; Indels 6; Gaps 5;

Qy      3 TGAACGCTGGCGGATGCTTACATGCAAGTGAACGCGACGAGTGTGCAATCT 62
Db      1 TGAACGCTGGCGGAGCTTACATGCAAGTGAACGCGAGTGTGCAATCT 59
Qy      63 GGTGCGAGTGGCGGACGCGGTGAGTAAATGCAATCGGAACGTAATCCAGAAAGAGGGGCTTAA 122
Db      60 GAATTC-AGCGGCGGACGCGGTGAGTAAATGCTTACGAAATCTGCTGTAGTGGGGAACAAC 118
Qy      123 GCATCGAAGATGTTGCTTAATCCGATATATCTTAAGAGAGAAAGACAGGGATCGAAGA 182
Db      119 GTCTCGAAGAGGACGCTAATACCGCATACGCTCTACGGAGAGAAAGACAGGGGACCTTGGG 178
Qy      183 CTTGCGCTTTTGGAGCGGCGGATGTCTGATTAGCTAGTTAGTGGGGTAAAGGCTTACCA 242
Db      179 CTTGGGCTATCAGATAGAGCTAGGCTAGGATAGCTAGTTAGTGGTAAATGGCTTACCA 238
Qy      243 AGGCGACATCATAGTATGTTGGTCTGAGAGAGACACACCACTGGGACCTGAGACAGGC 302
Db      239 AGGCGAGATCCGTAACCTGCTGAGAGAGATCACTCACTGGAACCTGAGACAGCT 298
Qy      303 CCAAGACTCTACGGGAGGACAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATCCA 362
Db      299 CCAAGACTCTACGGGAGGACAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATCCA 358

```

Db 299 CCAAGACTCTACGGGAGGACAGTGGGAAATATTGGACAATGGGCGAAAGCCTGATCCA 358  
Qy 363 GCAATGCCGGGTGAGTGAAGAAGCCCTTCGGGTTGTAACTCTTTCAAGTCGAAGAATA 422  
Db 359 GCCATGCCGGGTGAGTGAAGAAGCCCTTCGGGTTGTAACTCTTTCAAGTGGAGGAG 418  
Qy 423 AGGTTACGGTAAATTAATCGTGAACCATGACGGTATCGACAGAGAAACACCGGCTAACTA 482  
Db 419 GGCATTAACTTAATTAATCGTGAACCATGACGGTATCGACAGAGAAATTAACACCGGCTAATC 478  
Qy 483 CGTGCACGACAGCCGGTAAATCGTAGGGTGAAGCCCTTAATCGGAATTAATCGGGCTAA 542  
Db 479 TGTGCACGACAGCCGGTAAATCGTAGGGTGAAGCCCTTAATCGGAATTAATCGGGCTAA 538  
Qy 543 AGGGTGGCAGAGCCGGTAAATCGTAGGGTGAAGCCCTTAATCGGGCTTAACCTGGGAATTG 602  
Db 539 AGGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 598  
Qy 603 CGTTGAAATCTACAAAGCTAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 662  
Db 599 CATTCAAAATCGACAGCTAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 658  
Qy 663 AAATGCGTAGAGATATGGAAGAAATCATGATGGCGAAGGACCTCTGGGTTAACTGA 722  
Db 659 AAATGCGTAGATATGGAAGAAACACAGTGGCGAAGGACCACTGGACTGATCTGA 718  
Qy 723 CGCTCATGCAAGAAAGGTGGGAGCAACAGATTAAGATCCCTGGTATCCACCGCCT 782  
Db 719 CACTGAGGTGGGAAAGGTGGGAGCAACAGATTAAGATCCCTGGTATCCACCGCCT 778  
Qy 783 AAACGATGTAACATAGTTGTTGGGCTTATTAAG-GCTTGGTAAAGAACTAAAGCGGTAA 841  
Db 779 AAACGATGTAACATAGCCGTTGGGAGCCTTGAAGCTCTTAATGGCGCAGCTAAAGCATTA 838  
Qy 842 GTTGAACCGCTGGGAGTACGCTGGCAAGTAAACTCAAGAGATTGAAGCGGAGCCG 901  
Db 839 GTTGAACCGCTGGGAGTACGCGCGCAAGGTTAAACTCAAGAGATTGAAGCGGAGCCG 898  
Qy 902 CACAAAGGCTGAGTATATGATGATTAATTCGATGCAAGCGCAAAACCTTAACCTT 961  
Db 899 CACAAAGGCTGAGTATATGATGATTAATTCGATGCAAGCGCAAAACCTTAACAGGCTT 958  
Qy 962 ACATGTAACGAAATTTTCTAAGATATGATATGATG-CTTCGGGAAAGCTTAACACAGGCTG 1020  
Db 959 ACATGTAACGAAATTTTCTAAGATATGATATGATGCTTCGGGAAACATTAAGACAGGCTG 1018  
Qy 1021 CATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
Db 1019 CATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1078  
Qy 1081 CTTGTCAATTAATGCTCAATCTT--GGTTGGGCACTTAATGAGCTGCGGCTGAACAAAC 1138  
Db 1079 CTTGTCAATTAATGCTCAACGCTAATGCTGAGGCACTTAAGAGACTGCGGCTGAACAAAC 1138  
Qy 1139 CGGAGGAAGGTGGGGAATGAGCTCAAGTCCCTCATGGCCCTTAATGAGGCTTACACAGT 1198  
Db 1139 CGGAGGAAGGTGGGGAATGAGCTCAAGTCCCTCATGAGCCCTTAATGAGGCTTACACAGT 1198  
Qy 1199 AATACATAGCGCGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAAGAAAGCG 1258  
Db 1199 GCTACATAGTGGTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAAGAAAGCG 1258  
Qy 1259 CGTGTGATGCTGGATGAGGATCTGCAACTGACTCCGTAAGTCCGGAATGCTAATGTAATC 1318  
Db 1259 GATGTGATGCTGGATGAGGATCTGCAACTGACTCCGTAAGTCCGGAATGCTAATGTAATC 1318  
Qy 1319 GCGGATCAGCATGTCGGGTGAATACGTTCCGGGCTTGTACACACCGCCGCTCAACCC 1378  
Db 1319 GCGGATCAGCATGTCGGGTGAATACGTTCCGGGCTTGTACACACCGCCGCTCAACCC 1378  
Qy 1379 ATGGAGTGGGTTTCAACAGAGCAGGTAATCTTAACGCTAAGAGAGGCGCTTGGCCAGCT 1438  
Db 1379 ATGGAGTGGGTTTCAACAGAGCAGGTAATCTTAACGCTTGGAGAGAGCGGTTTACAGCGGT 1438

Qy 1439 GAGATTCATGACTGGGCTG 1457  
Db 1439 GTGATTCATGACTGGGCTG 1457

RESULT 30  
US-09-737-297-4  
; Sequence 4, Application US/09737297  
; Patent No. 6887984  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Mark  
; APPLICANT: Griffiths, Allen  
; APPLICANT: Hill, Philip  
; APPLICANT: Laybourne-Parry, Johanna  
; APPLICANT: Mills, Sarah  
; TITLE OF INVENTION: Processes and Organisms for the Production of Antifreeze Proteins  
; FILE REFERENCE: F3247  
; CURRENT APPLICATION NUMBER: US/09/737,297  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: GB 9929696.4  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 4  
; LENGTH: 1481  
; TYPE: DNA  
; ORGANISM: Marinomonas communis  
; FEATURE:  
; NAME/KEY: Unsure  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (203)..(204)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (840)..(840)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (964)..(965)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (1142)..(1142)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (1182)..(1182)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (1185)..(1186)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (1449)..(1449)  
; OTHER INFORMATION: base identity unsure  
US-09-737-297-4

Query Match 69.5%; Score 1012.6; DB 3; Length 1481;  
Best Local Similarity 82.1%; Pred. No. 0;  
Matches 1193; Conservative 1; Mismatches 255; Indels 4; Gaps 3;

Qy 1 ATTTGAAGCTGGCGGATGCTTTTACATGCAATGCAAGTCGAACGGCAGCAGAGTGTTCGAT 60  
Db 29 ATTTGAAGCTGGCGGATGCTTTTACATGCAATGCAAGTCGAACGGCAGCAGAGTGTTCGAT 88  
Qy 61 CT--GGTGGAGTGGCGGAGCGGCTGAGTATGCTGCAATGCAATGCAATGCAATGCAATGCA 118  
Db 89 AGAATATGACGAGCGCGGAGCGGCTGAGTATGCTGCAATGCAATGCAATGCAATGCAATGCA 148  
Qy 119 TAAAGCATGGAAGATGCTTAATACCGCATTAATCTTAAGAGGAAGAGCGGAGTGA 178  
Db 149 CAACATGAGGAAGCGCATGCTTAATACCGCATTAATCTTAAGAGGAAGAGCGGAGTGA 208  
Qy 179 AAGACCTTGGCTTTTGAAGCGGCGGATGCTGATTAAGTGTGAGGAGTGAAGGCT 238

Db 209 CGAGCTTTGCTAATAGATGAGCTGCGTGAATTAGCTAAGTTGGTGGGTTAAAGCCT 268  
Qy 239 ACCAAGCGACGATACAGTAGTGGTCTGAGAGGACGACGCACTGAGGACTGAGACA 298  
Db 269 ACCAAGCGACGATCTCTACTGCTGCTGAGAGGATGATCAAGCCACACTGGGACTGAGACA 328  
Qy 299 CGAGCCGAGACTCTTACGAGAGGACGACGTGGGAAATTTTGGACATAGGCGCAAGCTTGA 358  
Db 329 CGAGCCGAGACTCTTACGAGAGGACGACGTGGGAAATTTTGGACATAGGCGCAAGCTTGA 388  
Qy 359 TCCAGCAATGCGCGTGAAGTGAAGAAGCCCTGGGTTGTAAGCTCTTTCAGTCGAGAA 418  
Db 389 TCCAGCAATGCGCGTGAAGTGAAGAAGCCCTGGGTTGTAAGCACTTTCAGAGTGAAG 448  
Qy 419 GAAAAGGTTACGGTAAATTAATCGTACCCATGACGATGCGACGAGAAAGCAACCGGCTA 478  
Db 449 GAAAGGCGTATAGTTAATCTCTGATGTTTGAAGCTTACCCGAGAAAGCAACCGGCTA 508  
Qy 479 ACTACGTCCAGCAGCCGCGGTAAATACGTAGGGTGCAGCGTTAATCGGAATTACTGAGC 538  
Db 509 ACTCTGTGCGCAGCAGCCGCGGTAAATACGAGGGTGCAGCGTTAATCGGAATTACTGAGC 568  
Qy 539 GTAAAGGTTGCGCAGCGGCTTGAATGTCAGATGTGAATCCCGGCTTAACTGAGGA 598  
Db 569 GTAAAGGCGCGTGGCGGTTGTTAAGTGGATGTGAATCCGAGGCTCAACCTTGA 628  
Qy 599 ATTGCGTTTGAACCTACAAAGCTAGAGTGTGCGCAGAGGAGTGAATTCATGATGAGC 658  
Db 629 ATGACACCCGATACCTGCGAGGCTAGAGTACGATAGAGGGGTGTGAATTTCTGTGAGC 688  
Qy 659 AGTAAATGCGTAGAGATATGAAAGACATGATGCGAGGACAGCTCTGAGTTAACA 718  
Db 689 GGTAAATGCGTTAGTATAGAAAGAACATCAGTGGCGAGAGGCAACCTGAGCCGATA 748  
Qy 719 CTGACGCTCATGACGAAAGCTGGGGAGCAACAGAGTTAGATACCTGGTAGCTCAGC 778  
Db 749 CTGACGCTGAGGTGAAAGGCTGGGGAGCAACAGATTAATGATACCTGTGATGTCACG 808  
Qy 779 CCCCTAAAGCATGTCAATAGTTGTTGGGCTTATTAAGCTTGGTAAAGCAACCTACGCGT 838  
Db 809 CCGTAATGATGTCTACTAGCCGTTGGGAGATATATTTCTTATAGTGGCGACACTAACGGA 868  
Qy 839 GAAGTGACCGCTGGGGAGTACGCTGCGCAAGATTTAAATCTAAAGAAATGACGGGAC 898  
Db 869 TAAGTAGACCGCTGGGGAGTACGCGCCGAAGTTAAATCAATGATTTGACCGGCGC 928  
Qy 899 CCGCAAGCGGCTGATTAATGTGATTAATTCGATGCAACCGGAAACCTTACTACTCC 958  
Db 929 CCGCAAGCGGCTGATGTGTGTTAATTCGAANNAACCGGAAACCTTACTACTC 988  
Qy 959 TTGACATGAGGAATTTCTAGAGATAGATTAGT-CTTGGGAAACGCTTAACAGGCTG 1017  
Db 989 TTGACATGAGGAATTTCTAGAGATAGATTAGT-CTTGGGAAACGCTTACTACTACTG 1048  
Qy 1018 CTGATGAGCTGTCTGACGTGCTGTGCTGAGATGTTGGGTTAAGTCCCGCAAGAGCGCA 1077  
Db 1049 CTGATGAGCTGTCTGACGTGCTGTGCTGAGATGTTGGGTTAAGTCCCGTAAAGAGCGCA 1108  
Qy 1078 ACCCTTGATTAATTTGCCATCA-TTTGGTTGGGCACTTTAATGAGACTGCGCGTGAACA 1136  
Db 1109 ACCCTTATCTTATTTTGGCAGACACTTGGGTGGAACCTTAAGAGAGACTGCGGTGAACA 1168  
Qy 1137 ACCGAGGAAGGTGGGAGTGAAGTGAAGTCTCATGCGCCTTATAGGTTAGGGCTTCAAC 1196  
Db 1169 ACCGAGGAAGGTGGGAGTGAAGTGAAGTCAATCATATGCGCTTAAAGAGTCAACAC 1228  
Qy 1197 GTAAATACATGCGCGTACAGAGGTTGCAACCCGCGAGGGGAGCTTAATCTCAAGAA 1256  
Db 1229 GTGCTACATGCGCGTACAGAGGAGGAACTGCGCGAGGGTAAAGCAAAATCCCAAAAG 1288  
Qy 1257 CCGCTGTGATGCGGATCGGATCTGCAACTGCACTCCCTGAAGTCGAAATGCTAGTAA 1316  
Db 1289 TACGTGTAGTCCGATGGAAGTGTGCACTGACCTCAATGAAGTGGAGATGCTAGTAA 1348

Qy 1317 TCGCGATGACATGTCGCGGTGAATACGTTCCCGGCTCTTGATACACACCGCCCTCACA 1376  
Db 1349 TCGTGAATGAGATGTCAGGTGAATACGTTCCCGGCTCTTGATACACACCGCCCTCACA 1408  
Qy 1377 CCATGGAGTGGGTTTACACAGAGGATGATCTAATCCGTAAGAGGGGCTTGCCAGC 1436  
Db 1409 CCATGGAGTGGGTTTCTCAGAGTATGCTAATCCTTCGGAATGGCGGTTACACG 1468  
Qy 1437 GTGAGATTCTGA 1449  
Db 1469 GAGTGTGATGA 1481

RESULT 31  
US-08-114-695A-6  
Sequence 6, Application US/08114695A  
Patent No. 5508193  
GENERAL INFORMATION:  
APPLICANT: Mandelbaum, Raphael T.  
APPLICANT: Mackett, Lawrence P.  
TITLE OF INVENTION: DEGRADATION OF S-TRIAZINES IN SOIL AND  
TITLE OF INVENTION: WATER  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCHMEGMAN, LUNDBERG & WOESSNER, P. A.  
STREET: 3500 IDS CENTER  
CITY: MINNEAPOLIS  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/114,695A  
FILING DATE: 31-AUG-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MUEITING, ANN M.  
REGISTRATION NUMBER: 33,977  
REFERENCE/DOCKET NUMBER: 600.268US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-339-0331  
TELEFAX: 612-339-3061  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1518 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: rRNA  
ORIGINAL SOURCE:  
ORGANISM: Pseudomonas aeruginosa  
US-08-114-695A-6

Query Match 68.9%; Score 1003.4; DB 2; Length 1518;  
Best Local Similarity 67.1%; Pred. No. 0;  
Matches 975; Conservative 233; Mismatches 238; Indels 6; Gaps 6;

Qy 1 ATTGAACGCTGGCGCA-TGCTTTACATCAAGTGAACGCAACGCAACGATGCTTGA 59  
Db 28 AATGAACGCTGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCA 86  
Qy 60 TCTGTGCGCACTGCGGACGCGGAGTAAATGATCGGAACGTAATCCGAAGAGGGGCT 119  
Db 87 CCGGATGTC-AGCGCGGACGCGGAGGATGATGATGATGATGATGATGATGATGATGAT 145  
Qy 120 AACGATGGAAGATGCTAATACCGGATTAATCTTAAGAGGAGAAAGCAAGGAGATGAA 179  
Db 146 AACGCGGGAACGCGGCGCTUAUACGCAUACGUCCTUGAGGAGAAAGGAGGAGUCCUC 205

OY		180	A G A C T T G C G C T T T T T G A C C G C C G A T G C T G A T T T A G C T A G T T G G T G G G G T T A A G G C C T A	239
Db		206	G G A C C U C A C G C U A U C N G A U G A G C C U A G G C C G A U U A G C T A G U T G G G G G U A A G G C C T A	265
OY		240	C C A A G G C G C A G T A T C A G T A G T T G T C T G A B A G G A C G A C A G C A C A C T G G G A C T G A G A C A C	299
Db		266	C C A A G G C G A C G A U C C G U A U C U G S U C U G A B A G G A C A U C A G U C A C A C U G G A C U G A B A C A C	325
OY		300	G G C C C A G A C T C C T A C G G G A G G C A G C A G T E G G G A A T T T T G A C A A T G G C G C A A G C T G A T	359
Db		326	G G U C C A G A C U C C U A C G G G A G G C A G C A G U G G G G A A U A U V D G A C A U O G G G A A A G C C M G A U	385
OY		360	C C A G C A A T G C C G C G T G A G T G A A G A A G G C C T T G C G G G T T A A A G C T T T C A G T C G A G A G	419
Db		386	C C A G C C A U B C C G C G U G U G A A B A B A G U C U C C G A U U G A A A A G C A C U U A A B U U G G A G C	445
OY		420	A A A A G T T A C G G T A A A T A A T C G T G A C C C A T G A C G G T A T C A C A A A A A G C A C G G C T T A	479
Db		446	A A G G C A C A G U A A U A A C C U G C U G U U U G A C G U U A C C A A C A G A A U A A G C A C C G C C U A A	505
OY		480	C T A C G T G C C A G A G C C G C G T A A T A C T A G G G T G C A A G C G T A A T C G A A T T A C T A G G G C	539
Db		506	C U U C C U G C C A G A G C C G C G U A A U A C G A A G S U G C G A G G G U A U A U C G A A U U A C U G G G G	565
OY		540	T A A A G G T G C G C A G C C G G C C T T G T A A G T C A G A T G T A A A T C C C G G G C T T A A C T T G G G A	599
Db		566	U A A A C G G C G U A G G U G U U C A G C A A G U A U G A U A U C C C G G G C U A C C U G G G A	625
OY		600	T T G C G T T G A A A C T A C A A A G C T A G A G T G T G C A G A G G A G G T G A A A T T C A T G T G T A G C A	659
Db		626	C U G C A U C C N A A A C U A C U A G C U A G A U A C G S U A G A G G G G G G A A U U C C U G U A V A G G	685
OY		660	G T G A A A T G C G T G A G A T A T G A A A G A A C T C A T T G C A A G C A G C C T C C T G G T T A A C A C	719
Db		686	G U G A A A U C G U A G A U A U A G A A G A A C A C A B U G G C G A A G G C A C C U G S A C U G A U A C	745
OY		720	T G A C G C T C A T G C A G A A A A C G T G G G G A G C A A C A G A G A T T A G T A C C T T G A T G T C A C G C	779
Db		746	U G A C A C U A G S U G C A A A A G C G U G G G A G C A A C A G A A U U A G A U A C C U G U A G U C C A G C	805
OY		780	C C T A A A C A T G C A C T A G T T G T T G G G C C T T A T T A G - G C T T G T A A C G A A G T A C G G T	838
Db		806	C G U A A A C A U U G C G A C U A G C C G U U Z G A U C C U G A G A U D U U A G U S C C G A C G U A A C G C G A	865
OY		839	G A A G T T G A C C G C C T G G G G A G T A C G G T C G C A G A T T A A A C T C A A A G A A T T G A C G G G A C	898
Db		866	U A A G U C G A C C G C C U G G G G A G U A C G G C C G G A A G U U A A A C U C A A U U G A A U U G A C G G G G C	925
OY		899	C C G C A A A G C G G T G G A T T A T G T G A T T A A T T T G A T G C A A C G C G A A A A A C T T A C C T A C C	958
Db		926	N N G C A A A G C G G U G A G C A U O G U G U U A A U U D G A A G C A A C G G A A A A C C U A C C U G G C	985
OY		959	T T G A C A T G A G C G A A T T T T C T A G A G A T A G A T T A A G T - C T T C G G A A A C G T A A C A G A G T G	1017
Db		986	U U G A C A U C U G A G A A C U U U C C A G A B A U G A U U G U G C C U C G G A A C U C A G A C A G A G U G	1044
OY		1018	C T G A C G G C T G T C G T C A G T C G T G T C G T G A G A T T T G S G T T A A G T C C G C A A C G A G C C A	1077
Db		1046	C U G C A U G C U G C G U C A C C U C G U C G U G A G A U A G U U G G U A A G U C C G U A A C G A G C C A	1105
OY		1078	A C C C T G T C A T T A A T T G C A T C A - T T T G S T G G G A C A C T T T A T G A A C T G C C G G A G A A	1133
Db		1106	A C C C U G U C C U A G U U A C C A G A C C C U C G S U G S G G A C U C U A A G G A G A C U G C C G G U A C A A	1165
OY		1137	A C C G A G A A G G T G G G A T G A C G T C A A G T C C T A C A T G A G C C T T A T T G G T A G G G C T T C A C A C	1198
Db		1166	A C C G A G G A A G U U G G A U G A C G U C A A G U C A U C A I A G C C C U A C G C A A G G C C U A C A C	1223
OY		1197	G T A A T A C A T G A C G C G T A C A H A G S G T T G C A A C C C G C A G G G G A G C T A A T T C A G A A A G	1255
Db		1226	G U G C U A C A U U G U C G U A C A A A G G U U C G A A G C G C A G U G G A C U A U C C A U A A A A	1283

QY	1257	CGCGTCGAGTCCCGGATCTCGAGTCTGCACTCCGACTCCGGAATCGGTACTGAA	1311
Db	1286	CCGAUCGGAAGUCCG3AUCGCGUCUGCAACUGCGUAAUCGGAACGCUAGUAA	1345
QY	1317	TCGGGATACAGCATGTGCGCGGTGAATAGTTCGCCGGGTCTTGTAACACGCCCGTCACA	1376
Db	1346	UCGUGAUGCAGAUUCUACCGUGAUAUAGUUCGCCGGGCTUUGUACAACGCCCGUGACA	1405
QY	1377	CCATGGAGTGGGTTTTCACCAAGAACAGTATGCTTAAACCTGTAAGAGGGCGCTTGCCAC	1438
Db	1406	CCAUGGAGUGUGUGUCUCACAGAAUGACUAGUUAACCGCAAGGGGACGUGUACACG	1465
QY	1437	GTGAGATTCAATG 1448	
Db	1466	GAGUGAUUCAUG 1477	
RESULT 32			
US-09-726-774-14			
Sequence 14, Application US/09726774			
Patent No. 6677153			
GENERAL INFORMATION:			
APPLICANT: Iversen, Patrick L.			
TITLE OF INVENTION: Antisense Antibacterial Method and			
TITLE OF INVENTION: Composition			
FILE REFERENCE: 0450-0032.30			
CURRENT APPLICATION NUMBER: US/09/726, 774			
CURRENT FILING DATE: 2000-11-29			
PRIOR APPLICATION NUMBER: US 60/166,150			
PRIOR FILING DATE: 1999-11-29			
NUMBER OF SEQ ID NOS: 139			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 14			
LENGTH: 1487			
TYPE: DNA			
ORGANISM: Shigella dysenteriae			
US-09-726-774-14			
Query Match 68.6%; Score 999.8; DB 3; Length 1487;			
Best Local Similarity 82.0%; Pred. No. 0;			
Matches 1200; Conservative 0; Mismatches 257; Indels 6; Gaps 4;			
QY	1	ATTGACGCTGGCGGCAATGCTTTACACATGCAGTGAACGG--CAGCAGCATGCTTG 57	
Db	9	ATTGAACGCTGGCGGCGAGCCTTAACATCATCAAGTCCAAACGGTAAAGAAAGCAGCTTGC 68	
QY	58	CATCTGTCGCGAGTGGCGGACGGGTGATGAATGATCGGAACGTATCCAAAGAGGGGG 117	
Db	69	TGTTTGCTGACGATGGCGGACGGGTGATGAATGCTTGGGAACCTGCTTGATGAGGGGG 128	
QY	118	GTAACGATCGAAAGATGTGCTAATATCCGATATACTCTTAAGAGGAAACAGGGGATCG 177	
Db	129	ATACTACTGGAACCGGTACTATATCCGATATAAGTCGCAAGACCAAAGAGGGGACCT 188	
QY	178	AAAGACCTTGCGCTTTTGGAGCGGCCGATGTCTGATTAGCTTAGTTGGTGGGTAAAGGCC 237	
Db	189	TCGGGCTCTTTGGCATGTGCTGATGTGCCCAGATGGGATTAGTAGGTGGGTAAACGGCT 248	
QY	238	TACCAAGCGCAATCATAGTATGTTGTTGTGAGAGACGACACGACCATCGGACTGAGAC 297	
Db	249	CACCTTAGCGCAATCTCTAGCTGTGTGTGAGAGGATGACACGACCATCTGAACTGAGAC 308	
QY	298	ACGGCCAGACTCCTTACGGAGAGGACGACATGGGGAAATTTTGGACAAATGGGCGCAAGCTG 357	
Db	309	ACGGTCCAGACTCTTACGGAGAGGACGACATGGGGAAATTTTGGACAAATGGGCGCAAGCTG 368	
QY	358	ATCCAGCAATGCGCGTGAATGAAGAGGCTTTCGGTTGTAAAGCTCTTTCAATCGAGA 417	
Db	369	ATGAGGCAATGCGCGGTGATGAAGAGGCTTTCGGTTGTAAAGCTCTTTCAATCGAGA 428	
QY	418	AGAAAGGTTACGGTAATATATGTGACCCATGACGATATGACAGAAAGACCGGCT 477	
Db	429	GGAAGGAGTAAGTTATATCTTTTGCTCAATTGACGTTAACCGGACAGAAAGACCGGCT 488	

[illegible]

Patent No. 6949362					
GENERAL INFORMATION:					
APPLICANT: Tomb, Jean-Francois					
APPLICANT: Bramucci, Michael G.					
APPLICANT: Cheng, Qiong					
APPLICANT: Kostichka, Kristy N.					
TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors					
FILE REFERENCE: CL1709 US NA					
CURRENT APPLICATION NUMBER: US/10/007,527A					
CURRENT FILING DATE: 2001-12-05					
PRIOR APPLICATION NUMBER: 60/254,868					
PRIOR FILING DATE: 2000-12-12					
NUMBER OF SEQ ID NOS: 30					
SOFTWARE: Microsoft Office 97					
SEQ ID NO 12					
LENGTH: 1424					
TYPE: DNA					
ORGANISM: Rhodococcus AN12					
US-10-007-527A-12					
Query Match	68.6%	Score 999.6	DB 3	Length 1424	
Best Local Similarity	82.7%	Pred. No. 0			
Matches 1168	Conservative	0	Mismatches 224	Indels 4	Gaps 3
QY	65	TGGCGAGTGGCGGACGGGTGAGTAATGCATCGAACGTATCCAGAAAGAGGGGGGTAACGC	124		
DB	29	TTGAGAGCGCGCGACGGGTGAGTAATGCATCGAACGTATCCAGAAAGAGGGGGGTAACGT	88		
QY	125	ATCGAAAGATGTCTAATACCGCATATCTTAAAGAGAAACAGAGGGGATCCGAAAGCC	184		
DB	89	TCGGAACGACGCTAATACCGCATATCTTAAAGAGAAACAGAGGGGATCCGAAAGCC	148		
QY	185	TTGGGCTTTTGAGACCGCGCATATGTCTAGTAGTTGTTGGGTAAAGGCTTACCAAG	244		
DB	149	TTGGGCTATACAGATGAGGCTTAGTGTGGATTAGCTAGTTGGTAGTAATGGCTTACCAAG	208		
QY	245	GCGACGATCAGTAAGTTGGTCTGAGAGGACGACCACTGGGACTGAGACAGCGGCC	304		
DB	209	GCGACGATCCGTAATCTGTCTGAGAGGATGATCACTGGAATCGAGACAGCGGTCC	268		
QY	305	AGACTCTTACGCGGAGCGACGAGTGGGGAATTTTGGACAATGGGGCCAGCCTGATCCAGC	364		
DB	269	AGACTCTTACGCGGAGCGACGAGTGGGGAATTTTGGACAATGGGGCCAGCCTGATCCAGC	328		
QY	365	AATGCCGGGTGAGTGAAGAAGGCCCTTGGGGTTGTAAGCTCTTTCAGTCCGAAAGAAAG	424		
DB	329	CATGCCGGGTGAGTGAAGAAGGCCCTTGGGGTTGTAAGCTCTTTCAGTCCGAAAGAAAG	388		
QY	425	GTTACGATAATATCGTGAACCATGACGATTCAGACAGAAAGAACCGGCTAACCTACG	484		
DB	389	CAGTTAATCTAATACGTGATGTTTGTGACGTTACGACAGAAATTAACAACCGGCTAACCTACG	448		
QY	485	TGCCAGACGCGCGGTAAATACGTAGGGTGCAAGGCTTAATCGGAATTACTGGGGCTTAAG	544		
DB	449	TGCCAGACGCGCGGTAAATACGTAGGGTGCAAGGCTTAATCGGAATTACTGGGGCTTAAG	508		
QY	545	GGTGGCGACGCGGCTTGAAGTCAAGATGTGAATCCCGGGCTTAACCTGGGAAATTGCG	604		
DB	509	CCGCGGTGAGTGTGTTTGAATGTGAATCCCGGGCTTAACTGGGAAATTGCG	568		
QY	605	TTTGAACATCAAAAGCTAGAGTGTGGACAGAGGAGTGGAAATTCATGTTAGCACTGAA	664		
DB	569	TTCAAAATCTGACTGACTAGTAGTATGTGTAGAGGGGTGTGAAATTTCTGTGTAGCGGTGA	628		
QY	665	ATGCGTAGATATGGAAGACATCGATGGCGAGCGACGCTCTCGGGTTAACTGTACG	724		
DB	629	ATGCGTAGATATGGAAGACACACAGTGGCGCAAGCGCACTCGGACTGTATCTGACA	688		
QY	725	CTCATGACGAAAGCTGTGGGAGCAAAACAGATTGATACCTCTGGATGTCACGCGCTTA	784		
DB	689	CTGAGGTGGAAGCGTGGGAGCAAAACAGATTGATACCTCTGGATGTCACGCGCTTA	748		
QY	785	AGGATGTCAACTAGTTGTTGGGCTTATTAGG-CTTGGTAAAGCTAACCGGTGAAGT	843		

Db	749	ACGATGTTCAACTAGCCGTTGGAGCCCTTGAGCTCTTTAGTGGCGCAGCTTAAGCATTAA	808
Qy	844	TGACCCGCTGGGAGTAAGTCGCAAGATTAACTCAAAGAAATTGA	903
Db	809	TGACCGGCTGGGAGTAACGGCCGCAAGGTTAAATCTCAAAATGAAATTGA	868
Qy	904	CAAGCGGTGGATTATGTGGATTAAATTCGATGCAACGCGAAAACTTACCTTAC	963
Db	869	CAAGCGGTGGAGCATGTGGTTAAATTTCAGAACGACGAAAGAACTTACAGGCTTTAC	928
Qy	964	ATGTAGCGAAATTTTCTAGAGATAGATTATGTG-CTTGCGGAAACGCTAACACAGTGTGCA	1022
Db	929	ATCCAAATAACTTTCTTAAGATAGATGTGTGCTTTCGGAAACATTGAACAGGTGCTCA	988
Qy	1023	TGGCTGTGTCGACGTCGTGTGTCGTGATGTTAGTGGGTTAAGTCCCGCAAGACGCAACCT	1082
Db	989	TGGCTGTGTCGACGTCGTGTGTCGTGATGTTAGTGGGTTAAGTCCCGTAAAGACGCAACCT	1048
Qy	1083	TGTCATTAAATGTCATCA-TTTTGTTGGGCACTTTAAATGAGACTGCGGTGACAAACCG	1140
Db	1049	TGTCCTTAGTTAACAGCAAGCTAATGTTGGGCACTTAAGAGAACTGCGGTGACAAACCG	1108
Qy	1141	GAGGAAGTTGGGAGTAGAGCTCAAGTCCCTCAATGGTAAAGGCTTCAACGTTAA	1200
Db	1109	GAGGAAGTTGGGAGTAGAGCTCAAGTCAATCAATGAGCCCTTAACGCTGGCTTCAACGTTGC	1168
Qy	1201	TACAATGCGCGTAGACAGAGGGTTGCGCAACCGCGAGGGGAGCTAATCTCAGAAAAGCGCG	1266
Db	1169	TACAATGCTCGGTACAGAGGGTTGCGCAACCGCGAGGGTGAAGCTAATCCAGAAAAGCGCA	1228
Qy	1261	TGTAATGTCGGATTCGAGTCTTGCAACTCGACTCGTGAAGTCGTAATGCTTAATGTC	1320
Db	1229	TGTAATGTCGGATTCGAGTCTTGCAACTCGACTCGTGAAGTCGTAATGCTTAATGTC	1288
Qy	1321	GGATCAGATGTCGGGAGTAATAGCTTCCGGGCTTTTATACACCGCGGTCACACCAT	1380
Db	1289	GAATCAGATGTCGGGAGTAATAGCTTCCGGGCTTTTATACACCGCGGTCACACCAT	1348
Qy	1381	GGGAGTGGGTTTACACAGAGCAGGTAGTCTTAACCGTAAGGAGGGCGCTTGCCACGGTGA	1440
Db	1349	GGGAGTGGGTTTACACAGAGGTAGTCTTAACCGTGGGAGGAGCGTTTACACAGGTGT	1408
Qy	1441	GATTCATGACTGGGGT	1456
Db	1409	GATTCATGACTGGGGT	1424
RESULT 34			
US-08-757-653-158			
; Sequence 158, Application US/08757653			
; Patent No. 5843669			
GENERAL INFORMATION:			
APPLICANT: Kaiser, Michael W.			
APPLICANT: Lyamichev, Victor I.			
APPLICANT: Lyamichev, Natasha			
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using			
TITLE OF INVENTION: Thermostable FEN-1 Endonucleases			
NUMBER OF SEQUENCES: 190			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Medien & Carroll, LLP			
STREET: 220 Montgomery Street, Suite 2200			
CITY: San Francisco			
STATE: California			
COUNTRY: United States Of America			
ZIP: 94104			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/757,653			

[illegible]

777 CGCCCTAAACGATGTCACCTAGTTGTTGGCCCTTATTAGCCTTG- TAAACGAAGCTAACG 835  
808 CGCCCTAAACGATGTCACCTAGTTGTTGGCCCTTATTAGCCTTG- TAAACGAAGCTAACG 867  
836 CGTGAAGTTGACCGCCCTGGGAGTACCGTGGCAAGATTAAACTCAAGAAATTTGACGGG 895  
868 CGTTAAGTCACCGCCCTGGGAGTACCGTGGCAAGATTAAACTCAAGAAATTTGACGGG 927  
896 GACCCCGCAAGCGGCTGATTTATGATGATTAATTCGATGCAAGCGGAAACCTTACTTA 955  
928 GGGCCCGCAAGCGGCTGATTTATGATGATTAATTCGATGCAAGCGGAAACCTTACTTA 987  
956 CCCTTGCATGATGCAAGTTTCTAGAGAT-AGATTAGTCTTCCGGGACGCTTAACAAG 1014  
988 GTCTTGACATCCACGGAAGTTTTCAGAGATGAGATGATGCTTCCGGGACGCTTAACAAG 1047  
1015 GTGCTGATGATGCTGCTGATGCTGCTGATGATGCTGATGATGCTGATGATGCTGATG 1074  
1048 GTGCTGATGATGCTGCTGATGCTGCTGATGATGCTGATGATGCTGATGATGCTGATG 1107  
1075 GCAACCTTGTGATTAATGCTGATC-ATTGGTTGGGCACTTTATGATGATGATGATGATG 1133  
1108 GCAACCTTGTGATTAATGCTGATC-ATTGGTTGGGCACTTTATGATGATGATGATGATG 1167  
1134 CAACCGGAGAGATGAGGAGTGAAGTCAAGTCTTCAATGCTTATGATGATGATGATGATG 1193  
1168 TAAATGAGAGAGATGAGGAGTGAAGTCAAGTCTTCAATGCTTATGATGATGATGATGATG 1227  
1194 CACGTAATACAAATGCGCGTACAGAGGTTGCCAACCCGCGAGGAGGAGTCAATCTCAGA 1253  
1228 CAGGTGCTACAAATGCGCGTACAGAGGTTGCCAACCCGCGAGGAGGAGTCAATCTCAGA 1287  
1254 AAGCGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1313  
1288 AAGCGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1347  
1314 TAATCGGAGATCAGATGCTGCGGATGATGATGCTTCCGGGCTTGTACACACCGCCCTC 1373  
1348 TAATCGGAGATCAGATGCTGCGGATGATGATGCTTCCGGGCTTGTACACACCGCCCTC 1407  
1374 ACACCATGAGAGTGGGTTTCAACGAAGCAGTACTTAACCGTAAAGAGGCGCTTCC 1433  
1408 ACACCATGAGAGTGGGTTTCAACGAAGCAGTACTTAACCGTAAAGAGGCGCTTCC 1467  
1434 ACGGTGATTCATGATGAGGCTG 1457  
1468 ACTTGTGATTCATGATGAGGCTG 1491

RESULT 35  
US-09-465-355-2  
Sequence 2, Application US/09465355  
Patent No. 6316194  
GENERAL INFORMATION:  
APPLICANT: Karm, Jonathan  
APPLICANT: Knowles, David  
APPLICANT: Mutchie, Alastair  
APPLICANT: Lenz, Georg  
TITLE OF INVENTION: Methods and Kits for Discovery of RNA-Binding Antimicrobials  
FILE REFERENCE: 22620/1150 (formerly 3950/85276)  
CURRENT APPLICATION NUMBER: US/09/465,355  
CURRENT FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: US 09/325,601  
PRIOR FILING DATE: 1999-06-03  
PRIOR APPLICATION NUMBER: GB 9812196.5  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: GB 9904790.4  
PRIOR FILING DATE: 1999-03-02  
PRIOR APPLICATION NUMBER: US 60/122,439  
PRIOR FILING DATE: 1999-03-02  
PRIOR APPLICATION NUMBER: US 60/088,241  
PRIOR FILING DATE: 1998-06-05

NUMBER OF SEQ ID NOS: 37  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 1542  
TYPE: RNA  
ORGANISM: Escherichia coli  
US-09-465-355-2  
Query Match 68.4%; Score 996.8; DB 3; Length 1542;  
Best Local Similarity 66.3%; Pred. No. 0;  
Matches 971; Conservative 234; Mismatches 252; Indels 7; Gaps 5;  
1 ATTGAAGCTGCGGCGATGTTTACATGCAAGTGAAGGAGGACG--GATGCTTGC 58  
28 AATGAAGCTGCGGCGATGTTTACATGCAAGTGAAGGAGGACG--GATGCTTGC 87  
59 ATC--TGGTGGAGTGGCGGAGCGGAGTGAATGATCGGAACTATCCAGAGAGGG 116  
88 UUCUUCUGUACGAGUGGCGGAGCGGAGTGAATGATCGGAACTATCCAGAGAGGG 147  
117 GGTAAACGATCGAAGATGCTTAATACCGCATTAATCTTAAGAGAGAAACGAGGATC 176  
148 GAUAACTAATCGAAGATGCTTAATACCGCATTAATCTTAAGAGAGAAACGAGGATC 207  
177 GAAAGACTTGGCTTTTGGAGCGGCGGATGCTGATGATGATGATGATGATGATGATG 236  
208 UUCGCGCTUUCGCAACGAGUGGCGGAGTGAATGATCGGAACTATCCAGAGAGGG 267  
237 CTACCAAGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 296  
268 UCACTUAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 327  
297 CACGCGCCAGATCTCTTACCGGAGGAGGAGTGGGAAATTTGCAATGAGCGGAGCT 356  
328 CAGGCTUAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 387  
357 GATCAGCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 416  
388 GATCAGCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 447  
417 AAGAAAGGTTAGCGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 476  
448 AAGAAAGGTTAGCGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 507  
477 TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 536  
508 UAACTUAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 567  
537 GCGTAAAGGTTAGCGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 596  
568 GCGTAAAGGTTAGCGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 627  
597 GAATGCGTTTGAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 656  
628 GAACGCAATGCGTAAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 687  
657 GCGTAAAGGTTAGCGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 716  
688 GCGTAAAGGTTAGCGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 747  
717 CACTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 776  
748 GATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 807  
777 GCGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 835  
808 GCGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 867  
836 GGTGAAGTTGACCGCCCTGGGAGTACCGTGGCAAGATTAAACTCAAGAAATTTGACGGG 895  
868 CGTGAAGTTGACCGCCCTGGGAGTACCGTGGCAAGATTAAACTCAAGAAATTTGACGGG 927  
896 GACCCCGCAAGCGGCTGATTTATGATGATTAATTCGATGCAAGCGGAAACCTTACTTA 955



Db	928	GGCCCCGCAAAAGCGGUGGAGCGAUGGUGUUAUUGAGUGCAACCGGAAAGAAACUUUACUUG	987
Qy	956	CCCTTGCATGTAGCGAAATTTTCTTAGAGAT^AGATTAGTCTTGGGAAACGCTAACACAG	1014
Db	988	GUCUUGACAUCCACCGGAAGUUUUCAGAAUUGAGAAUUGGUCUUCCGGGAACCGUGAGCAG	1047
Qy	1015	GTGCGCATGAGCTGTGCGTCAAGTCTGTGCGTAGTGTGAGATGTTGGAGTTAAAGCCCGCAACGAGC	1074
Db	1048	GUGCUGCAUUGGCUUGGUCUACAGCUCUGGUGUUBAAUUGGUGUAAUUGUCCCGCAACGAGC	1107
Qy	1075	GCAACCCCTTGTCAATTAATTTGCCATC^ATTGSGTGGGACATTTAATGAGACTGCCGATGA	1133
Db	1108	GCAACCCUUUACUUUUGUUGGUCACGCGGCCGGCCGGGAACUAAAGGAGACUGCCACUGA	1167
Qy	1134	CAAAACCGGAGGAAGGTGGGGATGAGTCGTCAAGTCCATAGGCCCTTAATGGATAGGGCTTCA	1193
Db	1168	UAAACUGGAGGAAGGUGGAGUAGAGUUAAGUCANCAUAGGCCUUAACAACAGGAGCUCUA	1227
Qy	1194	CACGTAATACATGGCGCGCTACAGAGGGGTTGGCAACCCGCGAGGGGGAGCTAATCTCGA	1253
Db	1228	CACGUGCUACAAUUGCGCGCAUACAAGAAAGGACUUGCGAGAGCAAGCCGACCUCUUA	1287
Qy	1254	AAGCGCGTGTAGTATCGCGGATGGAGATCTGCAACTCTCGGTGAAGTCCGAAATGCTGAG	1313
Db	1288	AAGUGCGUGCGAGUUGCCCGGAUUGGAGUUCUGCAACUUGACUUGAUGAAUUGCGAUAUUGCUAAG	1347
Qy	1314	TAATGCGGAGATCAGCATGTGCGGCGGAGTAACGTTCCCGGAGTCTTTATACACCGGCCGCTG	1373
Db	1348	UAAUUGUGAUAUCAGAAUUGCCACAGGUAUAUAGUUUCCCGGCGCUUUGUACAACCGCCGUC	1407
Qy	1374	AACCCATGGAGTGGGTTTCAACAGAGCAAGTAATCTTAACCGTAAGGAGGCGCTTGCC	1433
Db	1408	AACAUGGGAAGUGGUGUACAATAAGAAUAGUAGUACUUAACUUUGGAGGAGGCGCUUAC	1467
Qy	1434	ACGGTGAATTCATGACTCGGGGTG	1457
Db	1468	ACUUUGUAUCAUGACUGGGGUG	1491

RESULT 36  
US-08-520-946-158  
Sequence 158, Application US/08520946  
Patent No. 6372424  
GENERAL INFORMATION:  
APPLICANT: BROW, MARY ANN D.  
APPLICANT: LYAMICHEV, VICTOR I.  
APPLICANT: OLIVE, DAVID M.  
TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
TITLE OF INVENTION: PATHOGENS  
NUMBER OF SEQUENCES: 160  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MEDLEN & CARROLL  
STREET: 220 MONTGOMERY STREET, SUITE 2200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/520,946  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, PETER G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: FORS-01756  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410

Query Match	68.4%	Score 996.8	DB 3	Length 1542
Best Local Similarity	82.3%	Pred. No. 0		
Matches 1205	Conservative	0	Mismatches 252	Indels 7
			Gaps	5
1	ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAAACGGCAGACAG--GATGCTTGC	58		
28	ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAAACGGTAAACAGAACGCTTGC	87		
59	ATC--TGGTGGCGAGTGGCGGACGGGTGAGTAATGCTATCGAACCTATCCAGAGGGGG	116		
88	TTCTTTGCTGACGAGTGGCGGACGGGTGAGTAATGCTATCGAGAACTGCTGATGAGAGGG	147		
117	GGTAACGATGAAAGATGTGCTAATACCGCATATCTCTAAGAGAAAGCAGGGGATC	176		
148	GATATCTACTGAAACCGGTACTAATACCGCATATCTCTGCAAGACCAAGAGGGGGAGCC	207		
177	GAAAGACTTGGCGCTTTTGGAGCGGCGCATGTCTGATTTAGCTAGTTGGTGGGTAAAGCC	236		
208	TTCCGGGCTCTTGGCATGATGTGCGCCAGATGGGATTTAGCTAGTGGTGGGTAAAGCC	267		
237	CTACCAAGCGCAGCATGATGATGTTGCTGTGAGAGGACGACACACTGGGAGCTGAGA	296		
268	TCACCTTACGCGCATGCTCTTACGCTGCTGTGAGAGATTAACAGCCACACTGGAACCTGAG	327		
297	CACGCGCCAGACTCTTACGCGGAGCGACAGTGGGAAATTTTGAACAATGGCGCAAGCCT	356		
328	CACGGTCCAGACTCTTACGCGGAGCGACAGTGGGAAATTTTGAACAATGGCGCAAGCCT	387		
357	GATCCAGCAATGCGCGCTGTGAGTGAAGAGGCTTGGGTTGTAAGCTCTTCACTGACAG	416		
388	GATCAGCGCAATGCGCGCTGTGAGTGAAGAGGCTTGGGTTGTAAGCTCTTCACTGACAG	447		
417	AAGAAAGGTTACGCTTAATATGTCGACCATGACCGGTATGACAGAAAGAACCGGC	476		
448	AGGAAAGGTTACGCTTAATATGTCGACCATGACCGGTATGACAGAAAGAACCGGC	507		
477	TAACTACGTCGACGAGCGCGGTATATCGTAGGGTGCAGCGTTATCGAATTAATCTGG	536		
508	TAACTCGGTGCAAGAGCGCGGTATATCGTAGGGTGCAGCGTTATCGAATTAATCTGG	567		
537	GCGTAAAGGTCGCGAGCGCGCTTGTATGTCAGATGTGAAATCCCGGGGCTTAACTGG	596		
568	GCGTAAAGGTCGCGAGCGCGCTTGTATGTCAGATGTGAAATCCCGGGGCTTAACTGG	627		
597	GAATTTGGCTTGAACCTACAAAGCTAGAGTGTGCGAGAGGAGAGTGTGATTCATGTGT	656		
628	GAACTGATCTGATATCTGCGAAGCTTGAAGTCTCTGTAAGGGGGGTGAAATTCAGGTGT	687		
657	GCACTGAAATCGTAGAGATATGGAAGAACATCGATGCGAAGGACGCTCTCTGGATTAA	716		
688	GCGGTGAAATCGTAGAGATATGGAAGAACATCGATGCGAAGGACGCTCTCTGGATTAA	747		
717	CAGTAGCGCTCATGACGAAAGCGTGGGAGCAAAACAGATTTAGTACCTCTGTGTGTCCA	776		
748	GACTGACGCTCAGAGTGCAGAAAGCGTGGGAGCAAAACAGATTTAGTACCTCTGTGTGTCCA	807		
777	CGCCCTAAACGATGTCACTAGTTGTTGGGCTTATTTAGGCTTGG--TAAAGAAAGCTAACG	835		
808	CGCCCTAAACGATGTCACTAGTTGTTGGGCTTATTTAGGCTTGG--TAAAGAAAGCTAACG	867		
836	CGTGAAGTGAACCGCTGGGAGTATCGGTCCGAAGATTAAACTCAAGAAATTTGACGGG	895		
868	CGTTAAGTGAACCGCTGGGAGTATCGGTCCGAAGATTAAACTCAAGAAATTTGACGGG	927		







QY 1134 CAACCCGAGAGAGTGGGATGAAGTCAAGTCTCAAGCCCTTATGAGGCTTCA 1193  
| | | | |  
Db 1168 TAAATCGAGAGAGTGGGATGACGTCAAGTCAATGAGCCCTTACGACAGGGCTTACA 1227  
| | | | |  
QY 1194 CAGCTAATACAAATGCGCCGTCACAGAGGTTGCCAACCCCGAGAGGAGCTAATCTCAGA 1253  
| | | | |  
Db 1228 CAGGTCTACAAATGCGCCATACAAAGAGAGCGACTCGGAGAGCAAGCGACTTATTA 1287  
| | | | |  
QY 1254 AAGGCGCTCGTAAGTCGGGATCGAGTCTGCAACTCGACTCGTAAGTCCGAATCGCTAG 1313  
| | | | |  
Db 1288 AAGGCGCTCGTAAGTCCGATTCGAGTCTGCAACTCGACTCGTAAGTCCGAATCGCTAG 1347  
| | | | |  
QY 1314 TAAATCGGATACAGCATGTGCGGCTGAATAGTTCGCGGCTTGTACACACGCGCCGTC 1373  
| | | | |  
Db 1348 TAAATCGGATACAGATGCGCAGCGGTGAATAGTTCGCGGCGCTTGTACACACGCGCCGTC 1407  
| | | | |  
QY 1374 ACACCATGGAGATGAGGTTTACACGAGAGCAGGTAGTCTAACCGTAAGAGAGGCGCTTGGC 1433  
| | | | |  
Db 1408 ACACCATGGAGATGAGGTTTACAAAGAGAGTAGTCTTAACCTTGGAGAGGCGCTTAAAC 1467  
| | | | |  
QY 1434 ACGGTGAGATTATGATGACTGGGCTG 1457  
| | | | |  
Db 1468 ACTTGTGATTATGATGACTGGGCTG 1491  
| | | | |  
RESULT 39  
US-10-061-071-33  
Sequence 33, Application US/10061071  
Patent No. 6894156  
GENERAL INFORMATION:  
APPLICANT: E.I. DUPONT DE NEMOURS & COMPANY  
APPLICANT: HENDRICKSON, EMIN  
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS FOR THE IDENTIFICATION OF DECHLORINATING  
TITLE OF INVENTION: BACTERIA  
FILE REFERENCE: BCI002 US CIP  
CURRENT APPLICATION NUMBER: US/10/061,071  
CURRENT FILING DATE: 2002-01-29  
PRIOR APPLICATION NUMBER: US 60/129,511  
PRIOR FILING DATE: 1999-04-15  
PRIOR APPLICATION NUMBER: 60/129,511  
PRIOR FILING DATE: 1999-04-15  
NUMBER OF SEQ ID NOS: 103  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 33  
LENGTH: 1542  
TYPE: DNA  
ORGANISM: Escherichia coli  
US-10-061-071-33  
Query Match 68.4%; Score 996.8; DB 3; Length 1542;  
Beeb Local Similarity 82.3%; Pred. No. 0; Mismatches 252; Indels 7; Gaps 5;  
Matches 1205; Conservative 0;

QY 297 CAGGCCCCAGACTCTTAACGGGAGGCGAGAGTGGGAAATTTTGGACAATGGGCGCAAGCT 356  
| | | | |  
Db 328 CAGGCTCCAGACTCTTAACGGGAGGCGAGAGTGGGAAATTTTGGACAATGGGCGCAAGCT 387  
| | | | |  
QY 357 GATCCAGCAATGCGCGGTGAGTGAAGAGCCCTTCGGTTGTAAAGCTCTTTCAGTGCAG 416  
| | | | |  
Db 388 GATCCAGCAATGCGCGGTGAGTGAAGAGCCCTTCGGTTGTAAAGTACTTTCAGCGGG 447  
| | | | |  
QY 417 AAGAAAGGTTACGTTAATATCTGACCATGACGGTATGCAACAGAAAGCACCGGC 476  
| | | | |  
Db 448 AAGAAAGGTTAAGTTAATATCTTGTCTATGAGCTTACCGCGAAGAAACACCGGC 507  
| | | | |  
QY 477 TAACTAGTCCAGCAGCCCGGTAAATAGCTAGGGTGAAGCGTTAATCGGAATTAATG 536  
| | | | |  
Db 508 TAACTCCGTCCAGCAGCCCGGTAAATAGCTAGGGTGAAGCGTTAATCGGAATTAATG 567  
| | | | |  
QY 537 GCGTAAAGGTTGCGAGGCGGCTTGTAACTGATGATGTAATCCCGGCTTAACTG 596  
| | | | |  
Db 568 GCGTAAAGGTTGCGAGGCGGCTTGTAACTGATGATGTAATCCCGGCTTAACTG 627  
| | | | |  
QY 597 GAAATGCGTTGAAACTACAAAGCTAGAGTGTGCGAGAGGAGTGAATTTCCATGTGA 656  
| | | | |  
Db 628 GAACTGATCTGATTAATCTGCAAGCTTGAAGTCTGTAAGAGGAGGTGAATTTCCAGGTGA 687  
| | | | |  
QY 657 GCACTGAATGCGTGAAGATATGGAAGATGAGTGGGAGAGGAGGAGCTCTCGGTTAA 716  
| | | | |  
Db 688 GCGTGAATGCGTGAAGATCTGAGAGATTCGAGGATTCGAGAGGAGGAGGAGGAGGAGGAG 747  
| | | | |  
QY 717 CACTGACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAAGTACCTGTGATGCA 776  
| | | | |  
Db 748 GACTGACGCTCATGCGGAGAAAGCGTGGGAGCAAAACAGATTAAGTACCTGTGATGCA 807  
| | | | |  
QY 777 CGCCTTAAAGATGCAACTAGTGTGGGCTTATTAAGCTTGG-TAAGACCTAAG 835  
| | | | |  
Db 808 CGCCTTAAAGATGCAACTAGTGTGGGCTTATTAAGCTTGG-TAAGACCTAAG 867  
| | | | |  
QY 836 CGTGAAGTTGACCGCTCGGAGAGTACCGTGGCAAGTTAAACCTCAAGAGAAATGAGCGG 895  
| | | | |  
Db 868 CGTGAAGTTGACCGCTCGGAGAGTACCGTGGCAAGTTAAACCTCAAGAGAAATGAGCGG 927  
| | | | |  
QY 896 GACCCGCAAGAGCGGTGATATGATGATTAATTCATGCAACCGCAAAACCTTACTTA 955  
| | | | |  
Db 928 GCGCCGCAAGAGCGGTGATATGATGATTAATTCATGCAACCGCAAAACCTTACTTA 987  
| | | | |  
QY 956 CCTTGAATGATGCAATTTTCTAGAGAT-AGATTAGTCTTGGGAAAGCTTAACAG 1014  
| | | | |  
Db 988 GTCTTGACATTCAGAGAAAGTTTTCAGAGATGAGAAATGCTTGGGAAAGCTTAACTG 1047  
| | | | |  
QY 1015 GTGCTGATGAGTGTGCTGACGCTGCTGCTGAGATGTTGGGTTAATCCCGCAACGAGC 1074  
| | | | |  
Db 1048 GTGCTGATGAGTGTGCTGACGCTGCTGCTGAGATGTTGGGTTAATCCCGCAACGAGC 1107  
| | | | |  
QY 1075 GCAACCTTGCATTAATGCAATC-ATTGGTGTGGGCACTTAAATGAGATGCGCGTGA 1133  
| | | | |  
Db 1108 GCAACCTTGCATTAATGCAATC-ATTGGTGTGGGCACTTAAATGAGATGCGCGTGA 1167  
| | | | |  
QY 1134 CAACCCGAGAGAGTGGGATGAGCTCAAGTCTCAATGCGCTTATGAGTGGCTTCA 1193  
| | | | |  
Db 1168 TAAATCGAGAGAGTGGGATGAGCTCAATGCTCAATGCGCTTATGAGTGGCTTCA 1227  
| | | | |  
QY 1194 CAGCTAATACAAATGCGCGGTACAGAGGTTGCCAACCCCGAGAGGAGCTAATCTCAGA 1253  
| | | | |  
Db 1228 CAGGTCTACAAATGCGCCATACAAAGAGAGCGACTCGGAGAGCAAGCGACTTATTA 1287  
| | | | |  
QY 1254 AAGGCGCTCGTAAGTCCGATTCGAGTCTGCAACTCGACTCGTAAGTCCGAATCGCTAG 1313  
| | | | |  
Db 1288 AAGGCGCTCGTAAGTCCGATTCGAGTCTGCAACTCGACTCGTAAGTCCGAATCGCTAG 1347  
| | | | |  
QY 1314 TAAATCGGATACAGATGCTCGGTTGAATAGTTCGCGGCTTGTACACACGCGCCGTC 1373  
| | | | |  
Db 1348 TAAATCGGATACAGATGCTCGGTTGAATAGTTCGCGGCTTGTACACACGCGCCGTC 1407  
| | | | |



```
; SEQ ID NO 2
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Salmonella typhimurium
US-09-726-774-2

Query Match      68.1%; Score 991.8; DB 3; Length 1541;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1195; Conservative 0; Mismatches 262; Indels 6; Gaps 4;

QY 1 ATTGAAGCTTGGCGGCGATGCTTTACATGCAAGTCCGAAACGG---CAGCAAGGATGCTG 57
DB 28 ATTGAAGCTTGGCGGCGGCGCTTAACATGCAAGTCCGAAACGGTAAACGAAAGCACTGTC 87
QY 58 CATCTGGGCGGAGTGGCGGCGGAGTGAATGATGCAATGCAAGTATCCAGAAAGGAGG 117
DB 88 TCTTGTGATGAGTGGCGGCGGAGTGAATGATGCTGGGAAATGCTGATGAGAGGAGG 147
QY 118 GTAACGATCGAAAGATGTGCTAAATACCGCATATATCTTAAGAGAAAGAGAGGAGATCG 177
DB 148 ATATCTACTGAAACGGTGGCTAAATACCGCATATATGTCGAAAGACAAAGAGGAGGACT 207
QY 178 AAAGACCTTGGCGCTTTTGGAGCGGCGCATGCTGATTAAGTGAATGTTGGGTAAAGGCC 237
DB 208 TCGGCGCTCTTGGCATGCGATGTGTGCCAGATGGGATTAAGTGAAGTGGGATTAAGGCT 267
QY 238 TACCAAGGCGGAGCATGATGATGCTGAGAGGAGCGACGACCACTGGGAGTGAAGC 297
DB 268 CACCTAAGCGGAGCATGCTTACCTGATGAGAGGATGACAGCCACTGAAGCTGAAGC 327
QY 298 ACGGCGGAGCATCTCTCAACGAGGAGCGAGCATGAGGAGAAATTTGAGCAATGAGGCGCA 357
DB 328 ACGGTCCAGATCTCTCAACGAGGAGCGAGCATGAGGAGAAATTTGAGCAATGAGGCGCA 387
QY 358 ATCCAGCAATGCGCGGTGAGTGAAGAGGCGCTTGGGTTGTAAGCTCTTTCAATCGAGA 417
DB 388 ATGAGCGCATGCGCGGTGATGAGAGAGGCGCTTGGGTTGTAAGCTCTTTCAAGGAGGA 447
QY 418 AGAAAGGTTACGGTAATTAATCGTGAACCAATGACGATGACAGAGAAAGCACGGCT 477
DB 448 GGAAGGTGTGTGTTAATTAATCCGACGATTAAGCTTAACCGGAGAAAGAGCACGGCT 507
QY 478 AACTACGTGCGAGCAGCGCGGTAAATAGTAGGAGTGAAGCGTTAATCGAATTAATCGAG 537
DB 508 AACTCCGTGCGAGCAGCGCGGTAAATAGTAGGAGTGAAGCGTTAATCGAATTAATCGAG 567
QY 538 CGTAAAGGATGCGAGCGCGGTGTAAGTCAATGATGTAATCCCGGGCTTAACCTGGG 597
DB 568 CGTAAAGGCGAGCGCGCGGTGTTAAGTCAATGATGTAATCCCGGGCTTAACCTGGG 627
QY 598 AATTGCGTTTGAATACTACAAAGCTGAGTGGCGAGAGGAGTGAATTCATGATGATG 657
DB 628 AACTGCACTGTATCTGAGCAAGCTTGAAGTCTGTAGAGGAGGAGTGAATTCAGATGATG 687
QY 658 CAGTGAATGCGTAGAGATATGGAAGAAATCATGATGCGAGAGCGACTCTTGGGTAAAC 717
DB 688 CGGTAAATGCGTAGAGATCTGGAAGAAATCCGATGCGAGAGCGCGCGCTTGGAGAG 747
QY 718 ACTGACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGGTATGTCAC 777
DB 748 ACTGACGCTCATGAGTGGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGGTATGTCAC 807
QY 778 GCCCTTAACGATGCACTAGTGTGTTGGGCTTTATTAAGGCTTGG--TAAAGAGCTTAAGC 836
DB 808 GCCGTAACGATCTTACTTGAAGTGTGCTTGAAGCGTGGCTTCCGAGACTTAAGC 867
QY 837 GTGAAGTTGACCGCTTGGGAGATACGATGCAAGATTAAACTCAAGAGATTGAAGCGG 896
DB 868 GTTAAGTTAGAGTCTTGGGAGATACGCGCAAGATTAAACTCAAGAGATTGAAGCGG 927
QY 897 ACCCGCAAGCGGTGATTAATGTGATTAATTCAGTGCAGCGGAAACCTTACCTTAC 956
DB 928 GCCCGCAAGCGGTGATTAATGTGATTAATTCAGTGCAGCGGAAACCTTACCTTAC 987
```

```
QY 957 CTTGACATGTAGCGAATTTTCTAGAGAT-AGATTAGTGTGGGAAAGCTAACACAGG 1015
DB 988 TCTTGACATCACAGAACTTTCCAGAGATAGATGTGCTTGGGAACTGTGAGACAGG 1047
QY 1016 TGGTGCATGCTGTGCTGACCTGCTGTGCTGAGATGTGGGTTAAGTCCCGCAACGAGCG 1075
DB 1048 TGTGCAATGCTGTGCTGACCTGCTGTGCTGAGATGTGGGTTAAGTCCCGCAACGAGCG 1107
QY 1076 CAACCTTGTCTATTATTTGCATATC-ATTGTTGGGCACTTTAATGAGACTGCGCGGTGAC 1134
DB 1108 CAACCTTATCTTGTGTTGTCAGAGGCTCCGCGGGAATCAAAAGAGACTGCGCATGAT 1167
QY 1135 AAACCGAGAGAGTGGGAGTGAACGTAACTCCATAGGCGCTTAATGGGTAGGCTTAC 1194
DB 1168 AAATGAGAGAGTGGGAGTGAACGTAACTCAATGAGCTTACCAACAGGCTTAC 1227
QY 1195 ACGTAAATCAATGCGGCTACAGAGGTTGCCAACCCGAGAGGAGGAGCTAATCTCAGAA 1254
DB 1228 ACGTAAATCAATGCGGCTACAGAGGTTGCCAACCCGAGAGGAGGAGCTTAA 1287
QY 1255 ACGCGTGTGATGCTCGGATCGAGTCTGCAACTGCTGCTGAGTGGATGCTTAC 1314
DB 1288 AGTGTGTGATGCTCGGATGCTGCAACTGCTGCTGAGTGGATGCTTAC 1347
QY 1315 AATGCGGATCAATGCTCGGCTGATTAAGTCTCCGCTTGTATACACACCGGCTTCA 1374
DB 1348 AATGCGGATCAATGCTCGGCTGATTAAGTCTCCGCTTGTATACACACCGGCTTCA 1407
QY 1375 CACCATGAGAGTGGGTTTCAACAGAGGAGTGAATCAACGTAAGAGGAGGCTTGC 1434
DB 1408 CACCATGAGAGTGGGTTTCAACAGAGGAGTGAATCAACGTAAGAGGAGGCTTGC 1467
QY 1435 CGGTGATTCATGACTGAGGCTG 1457
DB 1468 CTTGTGATTCATGACTGAGGCTG 1490

RESULT 42
US-09-737-297-1
; Sequence 1, Application US/09737297
; Patent No. 6887984
; GENERAL INFORMATION:
; APPLICANT: Berry, Mark
; APPLICANT: Griffiths, Allen
; APPLICANT: Hill, Philip
; APPLICANT: Laybourne-Barry, Johanna
; APPLICANT: Mills, Sarah
; TITLE OF INVENTION: Processes and Organisms for the Production of Antifreeze Protein
; FILE REFERENCE: F3247
; CURRENT APPLICATION NUMBER: US/09/737,297
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: GB 9929696.4
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Marinomonas protea
US-09-737-297-1

Query Match      67.8%; Score 987.6; DB 3; Length 1486;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 1206; Conservative 0; Mismatches 249; Indels 9; Gaps 6;

QY 1 ATTGAACGCTGGCGGCGATGCTT-TACATGCAAGTCCGAAAGCGAGCATGCTTGA 59
DB 11 ATTGAACGCTGGCGGCGAGCTTAAACATGCAAGTCCGAAAGCGAGGAGCTTGTCT 70
QY 60 TCTGAGGCGAGTGGCGGAGCGGTGATGATGCAATGCAATGCAATGCAATGCAATGCAAT 119
DB 71 CTTGTCAGAGAGCGGCGGAGCGGTGATGATGCAATGCAATGCAATGCAATGCAATGCAAT 130
```







QY 355 CTGATCCAGCAATGCGCGGTGAGTGAAGAAGCCCTTGCGGTGTGAAGCTCTTCACTGCG 414  
|||  
Db 361 CTGATGAGCCATGCGCGGTGAGTGAAGAAGCCCTTGCGGTGTGAAGCTCTTCACTGCG 420  
415 AGAAGAAAAGGTTACGGTAATATCGTGACCCATCGGTATCGACAGAAAGACCG 474  
|||  
Db 421 TGAGAGAAAGTTGGTGGCTTAATAGCGTATCAATTTGACGTGATGACAGAAAGACCG 480  
475 GCTAATACGTGCGACGACCGCGGTATATCGTAGGGTGCAGAGCTTAATCGAATTAAT 534  
|||  
Db 481 GCTAATCCGTGCGACGACCGCGGTATATCGTAGGGTGCAGAGCTTAATCGAATTAAT 540  
535 GGGCGTTAAAGGGTCCGAGCGCGCTTGTAGTCAATGTGAATCCCGGGCTTAACCT 594  
|||  
Db 541 GGGCGTTAAAGCGCATGAGCGCGGTGTGTTAAAGCAAGATGTGAAGCCCGGGCTTAACCT 600  
595 GGGAAATTGCGTTGAACTCAAAAGCTAGAGTGTGCGAGGAGGTGAATTCATGTC 654  
|||  
Db 601 GGGAAACCGCATTTTGAAGCTGCGACGGCTAAGAGTCTTGAAGGGGGGTGAAGATTTCA 660  
655 TAGCAGTGAATGCGTAGAGATATGAAGAACATGATGCGGAAAGCAGCCTCTGCGGTT 714  
|||  
Db 661 TAGCGGTGAATGCGTAGAGATCTGAAGAAATACCGGTGCGAAAGCGCGCCCTGAGACA 720  
715 AACACTGACGCTCATGACGAAAAGCGTGGGAGGAAACAGAGTTAGTACCTGATGTC 774  
|||  
Db 721 AAGACTGACGCTCATGATGCAAGAAAGCTGGGAGGAAACAGAGTTAGTACCTGATGTC 780  
775 CACGCCCTTAACAGATGTCAAC-TAGTGTGCGGCTTATTAAGGCTTGTGAACAGAGCTAA 833  
|||  
Db 781 CAGCGCTTAACAGATGTCTAATTGAGAGTGTGATCACTAAGTGTGCTTTCGAGCTAA 840  
834 CCGGTGAAGTTGACCGCTTGGGAGTACGCTGCGAGATTTAAACCTAAGAAATTGACG 893  
|||  
Db 841 CCGGTGAAGTTGACCGCTTGGGAGTACGCTGCGAGATTTAAACCTAAGAAATTGACG 900  
894 GGGACCGGCAAGCGGTGAGATTAATGAGATTAATGATGCAACCGGAAACCTTAAC 953  
|||  
Db 901 GGGACCGGCAAGCGGTGAGATTAATGAGATTAATGATGCAACCGGAAACCTTAAC 960  
954 TACCTTTGATGATGAGCAATTTTCTAGAGATTAATTAATGATG-CTTCGGGAACGCTAAC 1012  
|||  
Db 961 TACTCTTGAATCAAGAAAGCGGAAGAAATTTGGTGTGCTTTCGGAACCTGAGAC 1020  
1013 AGGTGCTGATGCGTGTGCTGACGCTGCTGTGAGATGTTGGTTAAGTCCGCAACGA 1072  
|||  
Db 1021 AGGTGCTGATGCGTGTGCTGACGCTGCTGTGAGATGTTGGTTAAGTCCGCAACGA 1080  
1073 GGGCAACCTTGTATTAATTTGCCATG-ATTGTTGGGCACTTTAATGATGATGCGCG 1130  
|||  
Db 1081 GGGCAACCTTGTATTAATTTGCTTTCGAGGAGATATGTGGGAATCTCCAGGAGATGCGG 1140  
1131 TGAACAAACCGGAGAAAGTGGGATGACGTCAGTCAAGTCTCATGCGCTTATGAGTGGCT 1190  
|||  
Db 1141 TGAATTAACCGGAGAAAGTGGGAGGACGATCAATGATCAATGAGCTTTAAGAGTGGCT 1200  
1191 TCAACGTAATTAACATGCGCGTACAGAGGTTGCCAACCGCGGAGGGGAGCTAATCTC 1250  
|||  
Db 1201 ACACACGCTGTACATATGAGCATATACAGAGGAGCGGAGCGGAGGTTGAGCAATCCC 1260  
1251 AGAAAGCGCTTGTATGTCGGATTCGGATTCGGATTCGGATTCGGATTCGGATTCGG 1310  
|||  
Db 1261 AGAAAGCGCTTGTATGTCGGATTCGGATTCGGATTCGGATTCGGATTCGGATTCGG 1320  
1311 TAGTAATCGCGGATCAGCATGTGCGCGGTGAATAGTTCGCGGCTTGTACACACCGCC 1370  
|||  
Db 1321 TAGTAATCGCGGATCAGCATGTGCGCGGTGAATAGTTCGCGGCTTGTACACACCGCC 1380  
1371 GTCAACCATGAGAGTGGGTTTCAACGAAAGCAGAGTATGTTAACCTGTAAGAGGGGCTT 1430  
|||  
Db 1381 GTCAACCATGAGAGTGGGTTTCAACGAAAGCAGAGTATGTTAACCTTGGGAGGGGCTT 1440  
1431 GCCACGCTGAGATTCATGACTGGGCTG 1457

Db 1441 ACCACGCTGTGCTTCAATGACTGGGCTG 1467  
RESULT 45  
US-09-492-709A-89  
Sequence 89, Application US/09492709A  
Patent No. 6720139  
GENERAL INFORMATION:  
APPLICANT: Zykkind, Judith  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Trawick, John  
APPLICANT: Forsyth, R. Allyn  
APPLICANT: Froelich, Jamie M.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN  
FILE REFERENCE: ELITRA.001A  
CURRENT APPLICATION NUMBER: US/09/492,709A  
CURRENT FILING DATE: 2000-01-27  
NUMBER OF SEQ ID NOS: 485  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 89  
LENGTH: 1549  
TYPE: DNA  
ORGANISM: E. Coli  
US-09-492-709A-89  
Query Match 67.7%; Score 985.8; DB 3; Length 1549;  
Best Local Similarity 82.3%; Pred. No. 0;  
Matches 1205; Conservative 0; Mismatches 252; Indels 8; Gaps 6;  
QY 1 ATGAAAGCTGGGCGGATGCTTAAACATGCAAGTGAACGCGACGAGAT--GCTTGC 58  
|||  
Db 28 ATGAAAGCTGGGCGGATGCTTAAACATGCAAGTGAACGCGACGAGATGCTTGC 87  
59 ATCT--GGTGGAGAGTGGGCGGATGCTTAAACATGCAAGTGAACGCGACGAGAT 116  
|||  
Db 88 TGGCTTGGCTGACAGAGTGGGCGGATGCTTAAACATGCAAGTGAACGCGACGAGAT 147  
117 GGTAAAGCTGGGCGGATGCTTAAACATGCAAGTGAACGCGACGAGATGCTTGC 176  
|||  
Db 148 GATTAATCTGAAACGATGCTTAAACATGCAAGTGAACGCGACGAGATGCTTGC 207  
177 GAAAGCTGGGCGGATGCTTAAACATGCAAGTGAACGCGACGAGATGCTTGC 236  
|||  
Db 208 TGGGCGCTTGGCGGATGCTTAAACATGCAAGTGAACGCGACGAGATGCTTGC 267  
237 CTACCAAGCGGATGCTTAAACATGCAAGTGAACGCGACGAGATGCTTGC 296  
|||  
Db 268 TCACCAAGCGGATGCTTAAACATGCAAGTGAACGCGACGAGATGCTTGC 327  
297 CACGCCCAAGCTCTTAAACATGCAAGTGAACGCGACGAGATGCTTGC 356  
|||  
Db 328 CACGCCCAAGCTCTTAAACATGCAAGTGAACGCGACGAGATGCTTGC 387  
357 GATTCAGCAATGCGGCTGATGTAAGAGAGGCTTGGGTTGTAAGCTCTTCACTGAG 416  
|||  
Db 388 GATTCAGCAATGCGGCTGATGTAAGAGAGGCTTGGGTTGTAAGCTCTTCACTGAG 447  
417 AAGAAAGGTTACGTTAATATCGTACCCATGACGAGTGAACGAGTGAACGAGTGAAC 476  
|||  
Db 448 AAGAAAGGTTACGTTAATATCGTACCCATGACGAGTGAACGAGTGAACGAGTGAAC 507  
477 TAACTCGTGGCGAGCGCGGTTAATAGTGAAGGTTAAGGCTTAACTGAACTTCTG 536  
|||  
Db 508 TAACTCGTGGCGAGCGCGGTTAATAGTGAAGGTTAAGGCTTAACTGAACTTCTG 567  
537 GCGTAAAGGTTGCGAGCGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 595  
|||  
Db 568 GCGTAAAGGTTGCGAGCGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 627



QY	596	GGAAATGGCGTTGAAACCTCAACAAACCTAGAGTGTGAGCAGAGGAGAGTGAATTCATGTGCT	655
Db	628	GGAACTGCATCTGATATCTGGCAACCTTAGAGTCTCGTAGAGGGGGGTAGAAATTCAGGTGT	687
QY	656	AGCAGTGAATATGCGTAGAGATATGAGAAACATTCGATGGCGAGGCGAGCTCTGGGTTA	715
Db	688	AGCGGTGAATATGCGTAGAGATCTGAGAGAAATACCGGTGGCGAAGCGGCCCCCTGAGAGA	747
QY	716	ACACTGAACGCTCATGACAGAAAGGTGGGGGAGCAACAGATTTGATACCTCGTAGATCC	775
Db	748	AGACTGACGCTCAGGTGCGAAGAGGTGGGGAGCAACAGATTTGATACCTCGTAGATCC	807
QY	776	ACGCGCTTAAACGATGTCACATAGTTGTTGGGCTTATTAAGGCTTGG-TAAAGAGCTAAC	834
Db	808	ACGCGGTAAACGATGTGTGACTTTGGAGGTGTGTCCTTGAGGCGTGGCTTCGCGAGCTAAC	867
QY	835	GCGTGAAGTTGACCGGCTCGGGGAGTACGGTGCGAAGATTTAAACTCAAGAAATTGACGG	894
Db	868	GCGTTAATCTCACCGGCTCGGGGAGTACCGGCGCAAGGTAAACCTCAATGAATATGACGG	927
QY	895	GGAACCCGCAACGCGGTGGAATTATGTGATTTAATTCGATGCACGCGAAAACCTTACT	954
Db	928	GGGCCCGCAACGCGGTGGAGCATGTGTTTAATTCATTCGACGCGAAGAACTTACT	987
QY	955	AACCTTGCATGTAGCGAATTTTCTAGAGAT-AGATTAGTCTCTTGGGAACTTAACACA	1013
Db	988	GGTCTTGACATCCACGGAAGTTTTCAGAGATGAGAATGTGCTTCGGGAAACCGTAGACA	1047
QY	1014	GGTGTGATATGCGCTGTGCTGACGCTCGTGTGCTGAGATGTTGGGTTAAGTCCGGAACGAG	1073
Db	1048	GGTGTGATATGCGCTGTGCTGACGCTGTGTGTGAATATGTGGGTTAAGTCCCGCAACGAG	1107
QY	1074	CGCAACCTTGTCAATTAATTGCGATC-ATTTGGTTGGCACTTAAATGAGCTGCCGTG	1132
Db	1108	CGCAACCTTATCTTTGTGTGCAACGGTCCGCGCGGGAACTCAAGAGAGCTGCCAGTGT	1167
QY	1133	ACAAACCGGAGGAAGGTGGGATGACGTCAAGTCTCATGGCCCTTAATGGGTAAGGGCTTC	1192
Db	1168	ATTAACCTGAGGAAGGTGGGATGACGTCAAGTCAATGAGCCCTTACGACACAGGGCTTAC	1227
QY	1193	ACACGTAATACATATGGCGCGTAGACAGAGGGTTGCGAACCCCGGAGGGGAGCTAAATCTCAG	1252
Db	1228	ACACGTGTCTACATATGGCGCATACAAAGAAAGCGACCTCGGAGAGCAAGCGGACCTCAT	1287
QY	1253	AAAGCGCGCTCGTAGTCCGATCCGATCGGAGTCTGCAACTGCACTCCGTGAATCGGAATCGCTA	1312
Db	1288	AAAGTGCCTCGTAGTCCGATCGGATTTGGAGTCTGCAACTGCACTCATATAATCGGAATCGCTA	1347
QY	1313	GTAATTCGGGGATCAGCATGTGCGCGGTGAATACGTTCCGGGGCTCTGTACACACGCCCGGT	1372
Db	1348	GTAATTCGGGGATCAGAAATGCCACGAGTGAATACGTTCCGGGGCTCTGTACACACGCCCGGT	1407
QY	1373	CACACCATGGAGTGGGTTTCCAGAAAGCAGTAGTCTAACCGTAAAGGAGGGCGCTTGC	1432
Db	1408	CACACCATGGAGTGGGTTTCCAGAAAGAGTAGTAGCTTAACTTCGGGAGGGCGCTTAC	1467
QY	1433	CACGGTAGATTCAATGACTGGGGGTG 1457	
Db	1468	CACCTTGTGATTCATGACTGGGGGTG 1492	

RESULT 46  
US-09-492-709A-242  
! Sequence 242, Application US/09492709A

: APPLICANT: Zyskind, Judith  
 : APPLICANT: Ohlsen, Kari L.  
 : APPLICANT: Trawick, John  
 : APPLICANT: Forsyth, R. Allyn  
 : APPLICANT: Froelich, Jamie M.  
 : APPLICANT: Carr, Grant J.

```

1  APPLICANT: Yamamoto, Robert T.
2  APPLICANT: Xu, H. Howard
3  TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
4  TITLE OF INVENTION: ESCHERICHIA COLI
5  FILE REFERENCE: ELITRA, 001A
6  CURRENT APPLICATION NUMBER: US/09/492,709A
7  CURRENT FILING DATE: 2000-01-27
8  NUMBER OF SEQ ID NOS: 485
9  SOFTWARE: FastSeq for Windows Version 3.0
10 SEQ ID NO 242
11 LENGTH: 1549
12 TYPE: DNA
13 ORGANISM: E. Coli
14 US-09-492-709A-242

```

Query Match	Score 985.8;	DB 3;	Length 1549;
Best Local Similarity	62.3%;	Pred. No. 0;	
Match 1205;	Conservative	0;	Mismatches 25;
			Indels 8;
			Gaps 6;

QY	1	ATTGAAAGCGTGGCGGCAATGCTTTTACATATGCAATGCAACGGGACGACGAGT - GCCTGC	58
Db	28	ATTGAAAGCGTGGCGGCAATGCTTTTACATATGCAATGCAACGGGACGACGAGT	87
QY	59	ATCT - GGTGCGAGTGGCGGACGGGTGATATGTCATCGGAACGTATCCGAAGAGGGG	116
Db	88	TGCTTCGCTGCGAGTGGCGGACGGGTGATATATCTGGGAAGCTGCTCATGGAAGGGG	147
QY	117	GGTAAAGCATGGAAGAATGTCTTAATCCGATATATCTTAAGGAGGAACAGAGGAGATC	176
Db	148	GATTAATACTAGGAAACGGTACTATATCCGATATATGTGCAAGACCAAGAGGGGAC	207
QY	177	GAAGAACCCTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTATGTGGGTAAAGGC	236
Db	208	TTCCGGCTCTTTCGCATCGAATGTGCCAAGTGGATTAGCTGTGGTGGGTAAACGGC	267
QY	237	CTACCAAGGCGACGATCAGTAGTGTGTGAGAAGACGACCAACCACTGGGACTGAGA	296
Db	268	TCACCAAGGCGACGATCCCTAGTGTGTGAGAAGATGACCAAGCACACTGGAACTGAGA	327
QY	297	CACGGCCGACACTCTTACGGGAGCAGAGTGGGGAATTTTGACAAATGGCGCAAGCT	356
Db	328	CACGTCGACACTCTTACGGGAGCAGAGTGGGGAATTTTGACAAATGGCGCAAGCT	387
QY	357	GATCCAGCAATCCGCGTAGTGAAGAAGGACCTTCGGTGTAAAGCTCTTTCAGTCAG	416
Db	388	GATGAGCGATAGCCGCGTAGTGAAGAAGGACCTTCGGTGTAAAGTACTTTCAGCGGG	447
QY	417	AAGAAAGGTTACGGTAAATATATCGTGAACCCATGACGGTATCGACGAAGAAGCACCGC	476
Db	448	AGGAAGGAGTAAAGTTAATATCTTGTCTATGACGTTACCCGAGAAAGACACCGC	507
QY	477	TAACTACGTGCCAGCAGCCGCGGTAAATACGTAGGGTGCAGGCTTAATCGAATTACTGG	536
Db	508	TAACTCCGTGCCAGCAGCCGCGGTAAATACGTAGGGTGCAGGCTTAATCGAATTACTGG	567
QY	537	GCGTAAAGGGTGGCGAGCGCGCTTG - TAAGTCAATGTGAATCCCGGGCTTAAACCTG	595
Db	568	GCGTAAAGCGCACGAGCGGGGTGGTTAAGTCATAGTGAATCCCGGGCTTCAACCTG	627
QY	596	GGAATTTGCTTTGAAACTTACAAAGCTTAAGTGTGCAGAGGAGGTGGAAATTCATGTGT	655
Db	628	GGAACTGATCTGAATCTGGCAAGCTTAAGTGTCTTAAGAGGGGGGTAGAAATTCACAGTGT	687
QY	656	AGCAGTGAATATGCGTAGAGATATGGAAGAACATCGATGGCGAAGCAGCCTCTGGGTTA	715
Db	688	AGCGGTGAATATCGTAGAGATCTGGAAGGAATACCGGTGGCGAAGCGGCCCTCGAGACA	747
QY	716	ACACTGACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTGAATACCTCTGGTATGTC	775
Db	748	AGACTGACGCTCATGAGTGCAGAAAGCGTGGGAGCAAAACAGATTGAATACCTCTGGTATGTC	807
QY	776	AGCGCTTAACGATGTCAATAGTTGTGGGCTTAATTAAGGCTTGG - TAAAGAGCTAAC	834

```

Db      808  AAGCGCTAACAATGTCAGACTTGGAGAGTTGTGCCCTTGAAGCGGTGCTTCCGAGCTAAC  867
Qy      835  GCGTGAAGTTGAACCGCCTTGCGGGAGTACCGTGCAGAACTAATAAATCTCAAGAAATTGACGG  894
Db      868  GCGTTAAGTCGACCGCCTTGCGGGAGTACCGCGCGAAGGTTAAATCTCAATGAATTGACGG  927
Qy      895  GGACCCCGCAACGCGGTGGATTATGTGGAATTAATTGATGCAACGGGAAAACCTTAACCT  954
Db      928  GGACCCCGCAACGCGGTGGAGCAATGTGTGTTAATTGCATCAACGGAAACCTTAACCT  987
Qy      955  ACCCTTGACATGTAGCGAAATTTTCTAGAGAT-AGATTAGTCTTCGGGAAACGCTTAACA  1011
Db      988  GGTCTTGACATCCACGGAAGTTTTCAGAGATGAGAATGTGCTTCGGGAACCGTGAGACA  1044
Qy      1014  GGTGCTGCATGGCTGTGTCAGCTCGTGTGTGTAAGATGTTGGGTTAATGTCGCAACGAG  1077
Db      1048  GGTGCTGCATGGCTGTGTCAGCTCGTGTGTGTAAGATGTTGGGTTAATGTCGCAACGAG  1110
Qy      1074  GCGAACCCCTGTGCTAATTAATTGGCATC-ATTGGTGGGGCACTTAATGAGACTGCGGATG  1133
Db      1108  CGCAACCTTATCTTTGTTGGCCAGCGGTCCGCGGGAATCTAAGAGACTGCAAGTG  1166
Qy      1133  ACAAACCGAGAGAGGTGGGGATGACGTCAAGTCTCATGCGCCCTTAATGGGTAGGGCTTC  1199
Db      1168  ATTAACCTGGAGAGAGGTGGGGAATGACGTCAAGTCAATCATATGCGCCCTTAACGACGAGGCTAC  1222
Qy      1193  ACACTGAATATCAATGGCGCGGTACAGAGGGTTGCCAACCCGCGAGGGGAGACTAATCTCAG  1255
Db      1228  AACCTGTCTACAAATGGCGCATACMAAGAGAGGACCTCGCGAGAACCAACGGAACCTCAT  1288
Qy      1253  AAAGCGCGTCTGAATGTCGCGATCGGAGTCTTGCAACTGCACTCGCTGAAGTGGAAATCGCTA  1311
Db      1288  AAAGTGGCGTCTGAATGTCGCGATGGAATCTTGCAACTGCACTGAAGTGGAAATGCTA  1344
Qy      1313  GTAATCGCGGATCAGCATGTGCGGGTGAATACGTTCCCGGGTCTTGTACACACGCGCGT  1377
Db      1348  GTAATCGTGAATCAGAAATGCCACGCGTGAATACGTTCCCGGGCTTGTACACACGCGCGT  1400
Qy      1373  CACACCATGGGAGTGGGTTTCACAGAGACAGTAATCTTAACCGTGAAGAGGGCGCTTGC  1433
Db      1408  CACACCATGGGAGTGGGTTTCACAAAGAAATGAGTAACTTAACCTTCGGGAGGGCGCTTAC  1466
Qy      1433  CACGGTGAAGTTCATGACTGGGGTG  1457
Db      1468  CACTTGTGATTCATGACTGGGGTG  1492

RESULT 47
US-09-492-709A-402
; Sequence 402, Application US/09492709A
; Patent No. 6720139
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froselich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; FILE REFERENCE: ELITRA-001A
; CURRENT APPLICATION NUMBER: US/09/492,709A
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 402
; LENGTH: 1549
; TYPE: RNA
; ORGANISM: E. Coli
US-09-492-709A-402

```

Query Match	67.7%	Score 985.8	DB 3	Length 1549	
Beet Local Similarity	66.3%	Pred. No. 0			
Matches	971	Conservative 234	Mismatches 252	Indels 8	Gaps 6
QY	1	ATTGNAACCTTGCGCGCATGCTTTTACACATGCAAGTCGAAACGGCAGCACGAT--GCTTGC	58		
DB	28	AUUGAACCTCUGCGCGGACGCCUACACAUUGCAAGUCGAAACGGUAAACAGAAAGCAGUCUG	87		
QY	59	ATCT--GGTGCAGAGTGGCGGACCGGTGATGATGTCATCGGAACGTATCCAAAGAGGGG	116		
DB	88	UGCUCUGCUGACGAGUGGCGGACCGGUGAGUAGUCUGGGAAGCUGCCUUAUGGAGGGG	147		
QY	117	GGTAAACGATCGAAAGAATGTGCTTAATACCGCATATACTTTAAGAGAGAAACACAGGGATC	176		
DB	148	GAUAACTUACUGGAAAACGGUACCUAAUAACCGACUAAUUGCCGAAAGCMAAAAGGGGAGC	207		
QY	177	GAAAGACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTAAAGC	236		
DB	208	UUCGGGCTCUCUGCCAUUGGAGUUGGCCAGAUUGGGAUUAGUCUUGUGGGGUAACGGC	267		
QY	237	CTACCAAGGCGACGATCAGTATGTTGTGCTGAGAGGACGACCAAGCCACCTGGGACTGAGA	296		
DB	268	UCACCAAGCGCAGCAUCCUGACUGUGUCUGAGAGAUAGACAGCCACACUGGAACUGAGA	327		
QY	297	CACGGCCCAAGCTCTCTACGGGAGCGACAGATGGGGAAATTTTGGACAAATGGGCGAAAGCT	356		
DB	328	CACGGUCCAGACUCCUACGGGAGCGACGAGGGGGAUUAUGCAACAAUGGGCGCAAGCCU	387		
QY	357	GATCCAGCAATGCGCGCTGATGAAAGAGGCGCTTCGGGTTTAAAGCTCTTTCAGTCAG	416		
DB	388	GAUCCAGCGCAUCCGCGUGUUAUGAAGAAAGCCUUCGGUUGUUAAAGUACUUCACGGGGG	447		
QY	417	AAGAAAAGTTACGGTAAATATATCTGACCCATGACGGTATCGACAGAGAAACACCGGC	476		
DB	448	AGGAAGGAGUAAAGUUAUAUACUUCUGCUCAUUGACUUAACCCGACAGAAAGACACGGC	507		
QY	477	TAATCTAGCTGCACAGACCGCGGCTAATACGTAAGGTGCAACGCTTAATCGAATTACTCG	536		
DB	508	UAAUCUGGUCACACGCGCGGUAUACGGAAGGUGCCAAACGUUAUCCGAUAUACUGG	567		
QY	537	GCGTAAAGGGGCGCAGCGCGGCTTG--TAAATCAGATGTGAAATCCCGGCGCTTAAACCTG	595		
DB	568	GCGUAAAGCGCACGCAAGCGGGUGUGUUAAGUCAGAUUGUAUACCCCGGCTUCAACUG	627		
QY	596	GGAATTGCGTTGAAACTCAAAAGCTAGATGTGGCAGAGGAGGTGGAAATTCATGTGT	655		
DB	628	GGAACUGAUCUAGUAUCUGGCACUCUUGAGUCUGUAGAGGGGGGUAUAUUCAGAGUGU	687		
QY	656	AGCAGTGAATGCGTAGATATGGAAGAAACATGATGGCGAAGGCAAGCCTCTCGGTTA	715		
DB	688	AGCGGUAAUUGCCUAAAGAUUCUGGAAGAAUACCGGUGGCGCAAGCGGCCCCCUGAGCGA	747		
QY	716	ACACTGAACGCTCATGCAAGAAAGGTGGGGAGCAACAGATTGATTAACCTGGTAGTCC	775		
DB	748	AGACUAGACGCUCAAGUGGAAAGCGUGGGAGCAACAGAUUUAUACCCUGGUAAGC	807		
QY	776	ACGCGCTTAAACGATGCAACTGATTGTTGGGCTTATTAGGCTTG--TAAAGCAAGCTAAC	834		
DB	808	ACGCGCUAAACGUAUGUGACUUGGAGGUGUGGCCUUGAGGCGUGGUCUUCGAGACUAC	867		
QY	835	GCGTGAAGTTGACCGGCTGGGAGTACGGTCGCAAGATTAAAACTCAAAGAAATTGACGG	894		
DB	868	GCGUUAAGUCACCGGCTUGGGGUAUACGGCGCAAGGCUUAAACUCAAUAGUAUUGACGG	927		
QY	895	GGAACCGGCAAGGCGGTGAGTTATGTGGATTAAATGATGCAAGCGGAAAAACTTACCT	954		
DB	928	GGGGCGCGCAAGGCGUGAGCAUAGUGUUUAUUCUAGUAGACGCAAGAAACCUUACU	987		
QY	955	ACCTTGACATGTATGCAAAATTTTCTTAAGAT--AGATTAGTCTTCGGGAAACGCTAACACA	1013		
DB	988	GGUUCUUAACUACCGAAGAUUUUACAGAUAGAAUAGUCUUCUGGGAACCGUAGACACA	1047		
QY	1014	GGTGTCTCAGGTGTGTGTGCTGCTGCTGTGTGTGATGATTGGTTAAATCCCGCAAGAG	1073		

```
Db 1048 GGUUCUGCAUUGGCUUCGUCAGUCUGUGUUAUUGUGUGUUAUUGCCCAAGAG 1107
Qy 1074 CGAACCCCTTGTATTAATTGGCATC-ATTGGATTGGGCACTTTAATGAGACTGCCGATG 1132
Db 1108 CCGAACCCCUUAUCCUUUGUGCCAGCGUCGCGCCGGAACUUAAGAGAGACGGCCAGUG 1167
Qy 1133 ACAAAACGAGAAAGGTGGGATGACGTCAAGTCCCTCATGGCCCTTAATGGGTAGGGCTTC 1192
Db 1168 AUAUAACUGAGAGAAAGUGGGGAGUAGCUCAAGUUAUAGGCCUUAACGACGAGGCUAC 1227
Qy 1193 ACAAGTAATCAATAGCGCGCTACAGAGGCTTCCAAACCCCGAGGGGAGACTAATCTCAG 1252
Db 1228 ACACGUGCUCAUAGUGGCGCAUACAAAGAGAGAGCCACUCCCGGAGAGCAACGCGACU 1287
Qy 1253 AAAGCGGTGTAGTCCGGATCCGAGTCCGATCCGATCCGGAATGCGGATGCTA 1312
Db 1288 AAAGUGGUGUGUAGUCGGAUUGGAGUCUCGACUCCAGCUCAGUAGUGGAAUUGCUA 1347
Qy 1313 GTAATCGCGATAGCATGTTCGCGTGAATACGTTCCCGGGTCTTGTACAACCGCCGCT 1372
Db 1348 GUUAUCUGUAGUACGAUUGCCACGUGUAUACGUCGCCGGCCUUGUAACACCGCCGU 1407
Qy 1373 CACACATGAGGAGTGGGTTTCAACCAAGACAGTAGTCTTAACCTTAAGAGGGCGCTTC 1432
Db 1408 CACACCAUUGGAGUGGUGUCAAAGAAAGUAGUAGCUUAACUUCGCGAGGGCGCUAC 1467
Qy 1433 CACGCTGAGATTCTGACTGGGCTG 1457
Db 1468 CACUUGUAGUUAUGACUGGGGUG 1492
```

```
RESULT 48
US-09-934-868-81
; Sequence 81, Application US/09934868
; Patent No. 6689601
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenle, Andreas J
; TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229, 858
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 81
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: METHYLOMONAS SP.
; US-09-934-868-81

Query Match 67.4%; Score 982.4; DB 3; Length 1429;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 1198; Conservative 0; Mismatches 226; Indels 10; Gaps 7;
```

```
Qy 253 CAGTATTGCTGTAGAGAGACGACCACTGGGACTGAGACAGGCGCCAGACTCCT 312
Db 237 CCGTACTGTGTGAGAGATGATATGACCACTGGGACTGAGACAGGCGCCAGACTCCT 296
Qy 313 ACGGAGGAGAGAGTGGGGAATTTTGAACAATGGGCGCAAGCTGATCCAGCAATCCG 372
Db 297 ACGGAGGAGAGAGTGGGGAATTTTGAACAATGGGCGCAAGCTGATCCAGCAATCCG 356
Qy 373 GTGAGTGAAGAAAGGCTTGGGTTGTAAAGCTTTTCACTCGAGAGAAAGTTACGGT 432
Db 357 GTGTGAGAGAAAGGCTTGGGTTGTAAAGCTTTCAATGGGAGAAACCTATCCGT 416
Qy 433 AATAATCGTACCCATGACCGTATGACAGAAAGAGCAACCGCTAATACGCGCAGCA 492
Db 417 TAATACCCGGTAC- TGAATTATCCATACAAAGAGCAACCGCTAATCTCGTCCAGCA 475
Qy 493 GCGCGGTAAATACGTAGGGTGCAGAGGCTTAATCGAATTACTGGGCTAAAGGGTGC 552
Db 476 GCGCGGTAAATACGTAGGGTGCAGAGGCTTAATCGAATTACTGGGCTAAAGGGTGC 535
Qy 553 GCGCGCTTGTAAAGTCAATGTGAAATCCCGGCTTAACTTGGAAATTCCTTTGAAC 612
Db 536 GCGCGTTTTTAAAGTCAATGTGAAATCCCGGCTTAACTTGGAAATTCCTTTGAAC 595
Qy 613 TACAAAGCTAGAGTGTGCGAGAGGAGTGGAAATTCAGTGTAGAGTGAATGGGTAG 672
Db 596 TGGGGAACCTAGAGTGAAGTGAAGAGAGTGAATTTTCAAGTGTAGCGGTAAATGGGTAG 655
Qy 673 AGATATGGAAGAACATCGATGGGAGAGGAGCGCTCTGGGTTAACTGACGCTCATGCA 732
Db 656 AGATCTGAAGAGAACACAGTGGGAGAGGCGGCTCTGACCTCAACTGACGCTGAGGTA 715
Qy 733 CGAAAGGTGGGAGAGCAACAGAGATTAGATACCTGTAGTCAACGCTTAAACGATGTC 792
Db 716 CGAAAGGTGGGAGAGCAACAGAGATTAGATACCTGTAGTCAACGCTTAAACGATGTC 775
Qy 793 AACTAGTTGTTGGG- CCTTATTAGGCTTGTGAACGAAGCTTAAGCCGTGAAGTGAACGGCC 851
Db 776 AACTAACCGTTGGGTTCTTAAAGAACTTATGTGTGAGCTTAAGCTTAAATGTTAGCCGC 835
Qy 852 TGGGAGTACGGTCCGAAGATTAAACTCAAGGAATTGACGGGAGCCGCAACAGCGGT 911
Db 836 TGGGAGTACGGTCCGAAGCTTAAACTCAAGGAATTGACGGGAGCCGCAACAGCGGT 895
Qy 912 GGATTATGGAATTAATTCGATGCAACGCCAAAAAACCCTTAACTAACCTTGAATGACG 971
Db 896 GGAAGATGTGTTAATTCGATGCAACGCCAAGAACTTAACCTTGAATGACCTCGG 955
Qy 972 AATTTCTAGAGATTAGATTAGTG- CTTCGGGAACGTTAACACAGTGTCTGATGGCTGT 1030
Db 956 AACTTGTCAAGATGACTGTGTGCTTCGGGAACCGAGAGACAGTGTCTGATGGCTGT 1015
Qy 1031 GTCAAGCTGTGTCTGAGATGTTGGGTTAAGTCCCGCAAGAGCGCAACCTTATCTTA 1090
Db 1016 GTCAAGCTGTGTCTGAGATGTTGGGTTAAGTCCCGCAAGAGCGCAACCTTATCTTA 1075
Qy 1091 ATTGCCA--TCAATTGTTGGGCACTTTAAATGAGACTGCGGTGACAAACCGAGAGAG 1148
Db 1076 GTTCCAGCGGCTATGCGCGGAACCTTAAGGAACTGCGGCTATTAACCGAGAGAG 1135
Qy 1149 TGGGAGTACGTCGAATCTCTATGCGCTTAATGGGTAGGGCTTCAACAGCTAATCAATG 1208
Db 1136 TGGGAGTACGTCGAATCTCTATGCGGCTTAATGGGTAGGGCTTCAACAGCTAATG 1195
Qy 1209 CGCTTACAGAGGTTTGCACCCGCGAGGGGAGCTATCTCAGAAAGCGGCTGTAGTC 1268
Db 1196 TCGGTACAGAGGTTTGCACCTCGAGAGCGCAATCCAAAGAGCGGCTGTAGTC 1255
Qy 1269 CGGATCGGAGTCTGCAACTCGAC- TCCGTGAAGTGGGAATCGCTAATTCGCGGATCAG 1327
Db 1256 CGGATTCGAGTCTGCAACTCGACTTGCATGAAGTGGGAATGCTAATTCGCGGATCAG 1315
```

Qy 1328 **AATGTCGGGGGAATACGTTCCCGGGCTTTGTACACACCGCCGTCACACCATGGAAATG** 1387

Db 1316 **AATCCCGGGGTAACTCTTCCCGGCTTTGTACACACCGCCGTCACACCATGGAAATG** 1375

Qy 1388 **GGTTTACACAGACGAGTAGTCTTAAACCGTAAGAGAGGGCGCTTGGCCACGCGTAG** 1441

Db 1376 **GGTTGCAAAAGAGTAGTATTAACTTCCGAGAGGGCGCTTACCATCTTTGTG** 1429

```

RESULT 49
US-10-701-200-81
: Sequence 81, Application US/10701200
: Patent No. 6958322
: GENERAL INFORMATION:
: APPLICANT: Koffas, Matheos
: APPLICANT: Odom, James M
: APPLICANT: Schenzle, Andreas J
: TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
: FILE REFERENCE: CL1596 US NA
: CURRENT APPLICATION NUMBER: US/10/701,200
: CURRENT FILING DATE: 2003-11-04
: PRIOR APPLICATION NUMBER: US/09/934,868
: PRIOR FILING DATE: 2001-08-22
: PRIOR APPLICATION NUMBER: 60/229,858
: PRIOR FILING DATE: 2000-09-01
: NUMBER OF SEQ ID NOS: 81
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 81
: LENGTH: 1429
: TYPE: DNA
: ORGANISM: METHYLOMONAS SP.
US-10-701-200-81

```

Query Match	67.4%	Score 982.4;	DB 3;	Length 1429;
Best Local Similarity	83.5%;	Pred. No. 0;		
Matches 1198; Conservative	0;	Mismatches 226;	Indels 10;	Gaps 7;

Oy	13	CGGATGCTTTAACAATCAGTCAGAGCGGAGCAGCGATGCTTGCACTCTGATGGCAGT	72
Db	1	CGGATGCTTTAACAATCAGTCAGAGCGGAGCAGCGATGCTTGCACTCTGATGGCAGT	56
Oy	73	GGCGGACGGGTGAATATGCAATCGGAACGTATCCAAAGAGGGGGGTAAACGATCGAAAG	132
Db	57	GGCGGACGGGTGAATATGCAATCGGAACGTATCCAAAGAGGGGGGTAAACGATCGAAAG	116
Oy	133	ATGAGCTAATAACCGCATATACCTTAAGAGGAAAGCAGGGGATCGAAAGACTTGGCGCTT	192
Db	117	TCACGCTAATAACCGCATATACCTTACGAGAGGAAAGCCGGGGAACCTTCCGGGCTGGCGCTA	176
Oy	193	TTGAGCGGCGGATGCTGATTTAGCTAGATTGGTGGGTTAAAGCCTTACCAAGCGACGAT	252
Db	177	ATAGATGAGCTAATGCTCGAATTAGCTATTGGTGGGTTAAAGCCTTACCAAGCGACGAT	236
Oy	253	CAGTAGTTGGTCTGAGAGGACGACACGAGCACTAGGCACTAGACACAGGGCCAGACTCT	312
Db	237	CCGTAAGCTGGTCTGAGAGGATGATAGCACTAGGCACTAGACACAGGGCCAGACTCTCT	296
Oy	313	ACGGGAGGACAGTGGGGAAATTTTGGCAATGGGCGGCAAGCCTGATCAGCAATGCGGC	372
Db	297	ACGGGAGGACAGTGGGGAAATTTTGGCAATGGGCGGCAAGCCTGATCAGCAATACCGC	356
Oy	373	GTCAGTGAAGAGGCTTCGGGTTGTAAAGCTCTTTCAGTGGAGAAAGAAAGTTACCGT	432
Db	357	GTCGTCGAAGAGGCTGAGGGTGTGTAAAGCACTTTCATGTGAAAGGAACCTATCCGT	416
Oy	433	AAATATATGTAACCATATGACGGTATCGACAGAGAGACCGGGCTAATCAATGATGACAGA	492
Db	417	TAAATACCCGGTAGAC-TGACATTAACCATACAAGAGACCGGGCTAATCCGTGACAGA	475
Oy	493	GCCCGGCTAATACTGATGGGTGCAGACGCTTAATCGAATTAATCTGGGCGTAAAGGGTGGCA	552
Db	476	GCCCGGCTAATAACGAGGGGTGCAGACGCTTAAATCGGAATTAATCTGGGCGTAAAGGGTGGCA	535

OY	553	GGCGGCGCTTGAATCACTCAGATGTGAAATCCCGGGCTTAACTCTGGAAATTCGTTTGAAC	612
Db	536	GGCGGTTTTTTTAAGTCAAGATGTGAAACCTCGGGCTTAACTCTGGAACTCCTCATTTGATAC	595
OY	613	TACAAAGCTAGAGTGTGGCAGAGGGAGGTGGAAATTCATGTGTAGCAGTGAATTCGCTAG	672
Db	596	TGGGGAACCTAGAGTTGAGTGTAGAGGAGGTGGAAATTCAGGTGTAGCGGTAAATTCGCTAG	655
OY	673	AGATATGGAAGAAACATGATGGCGAAGGACGCGCTCTGGGTTTAAACATGACGCTCATGCA	732
Db	656	AGATCTTGAGGAACACCACTGGCGAAGGCGGCTCTCTGGACTCAAACTGACGCTGAGGTA	715
OY	733	CGAAAGCGTGGGGAGCAAAACAGATTAGATACCTGTAGTCCACGCGCTTAAACGATGTC	792
Db	716	CGAAAGCGTGGGATGCAAAACGAGATTAGATACCTGTAGTCCACGCGCGTAAACGATGTC	775
OY	793	AACATAATTGTTGGG-CCTTATTAGGCTTTGGTACGAAGCTTACGCGGTGAATTTGACCGCC	851
Db	776	AACATAACCGTTGGGTTCTTAAAGAACTTATGTGTGAGCTTACGTAATTTAAGTGAACGCC	835
OY	852	TGGGAGATGACGATCGCAAGATTAAACTCAAGGAATTGACGGGGACCCGACACGCGT	911
Db	836	TGGGAGATGACGCGCGCAAGGCTAATACTCAATGAAATTGACGGGGCGCCGACACGCGT	895
OY	912	GGATTATGTGATTAATTCATGTCACGCGAATAAATCTTAACTTACCTTGACATGTAGCG	971
Db	896	GGACACATGTGGTTTAATTCATGTCACGCGAAGAACTTAACTTACCTTGACATCTCCTGG	955
OY	972	AATTTTCTAGAGATAGATTATAGT-CCTCGGGAAACGCTTACACAGAGTCTGTCATGCGTGC	1030
Db	956	AACCTTGTCAGATGACTGTGTGCTTCCTCGGAAACGAGAACAGAGTCTGTCATGCGTGC	1015
OY	1031	GTCAGCTCGTGTCTGTAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGTCTTAA	1090
Db	1016	GTCAGCTCGTGTCTGTAGATGTTGGGTTAAGTCCCGTAAAGAGCGCAACCTTAACTCTTAA	1075
OY	1091	ATTGCCA-TCATTGTGTTGGGCACTTTAATAGACCTGCGGTGACAAACCGGAGGAGG	1148
Db	1076	GTTGCCAGCGGCTATGCGCGGGAACCTTACGAGAGACTGCGGTGATTAACCGGAGGAGG	1135
OY	1149	TGGGAGATGACGTCAAGTCTCATAGGCCCTTAATGGGTGAGGCTTCAACGTAATACATAGG	1208
Db	1136	TGGGAGATGACGTCAAGTCTCATAGGCCCTTAATGGGTGAGGCTTCAACGTAATACATAGG	1195
OY	1209	CGCGTACAGAGGGTTGCCAACCGCGAGGGGAGACTAATCTCAAGAAAGCGCGTGTATGC	1268
Db	1196	TCGGTACAGAGGGTTGGAACTCCGCAAGCCAGCCCAATCCAAAAAGCGATCTATGTC	1255
OY	1269	CGGATTCGAGGCTCTCAACTCGAC-TCGGTGAAGCGGAATTCGCTAGTAATTCGCGGATCAG	1327
Db	1256	CGGATTCGAGGCTCTCAACTCGACTTTCGATGAGTCGGAATTCGCTAGTAATTCGCGGATCAG	1315
OY	1328	CATGTGCGGTGAATTAACGTTCCCGGATCTTGTACACACGCGCCGTCACACATAGGAGTG	1387
Db	1316	AATGCGCGGTGAATTAACGTTCCCGGCTCTGTACACACCGCCGTCACACATAGGAGTG	1375
OY	1388	GGTTTACACGAACAGATAGTCTTAACCGTAAGAGAGGCGCTTGCCACGGGTAG	1441
Db	1376	GGTTGCAAAAGAAAGTATGATGTTTAACCTTCGGAGGGGCGCTTACCACTTTGTG	1429

```

RESULT 50
US-09-596-002-41/C
; Sequence 41, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MOXAHELIA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16

```

PRIOR APPLICATION NUMBER: 60/140,121  
 PRIOR FILING DATE: 1999-06-18  
 NUMBER OF SEQ ID NOS: 41  
 SOFTWARE: PERL Program  
 SEQ ID NO 41  
 LENGTH: 269223  
 TYPE: DNA  
 ORGANISM: Moraxella catarrhalis  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Incyte template ID No. 6632636 41  
 PUBLICATION INFORMATION:  
 US-09-596-002-41

Query Match 67.2%; Score 979; DB 3; Length 269223;  
 Best Local Similarity 82.5%; Pred. No. 0;  
 Matches 1209; Conservative 0; Mismatches 240; Indels 16; Gaps 7;

1 ATTGAACGCTGGCGGCGCTTTACATGCAAGTCGACGCGACGCGATGCTTGCAT 60  
 92966 ATTGAACGCTGGCGGCGCTTTACATGCAAGTCGACGCGATGCTTGCAT 92908  
 61 CTGGTGGCGAGTGGCGGCGGCTGATGATGCAATCGAATCGAATCGAATCGAAT 120  
 92907 CTGATCTAGTGGCGGCGGCTGATGATGCAATCGAATCGAATCGAATCGAAT 92848  
 121 ACCGATCGAAGATGCTGAATACCGGATTAATCTTAAGGAGGAGGAGGATCGAAA 180  
 92847 ACTTGGGAGAACCCGACCTTAATACCGGATTAATCTTAAGGAGGAGGATCGAAA 92793  
 181 GACCTTGGCTTTTGGAGCGCGCGATGCTGAATGATGCTTGGGCTTAAAGGCTTAC 240  
 92792 TAGTCTCGCTATTAGTAGAGCTTAAGTGGATTAAGTGGGCTTAAAGGCTTAC 92733  
 241 CAAGGCGACATCATGATGCTTGGTGTGAGAGGACGACGACGACGACGACGACG 300  
 92732 CAAGGCGACATCATGATGCTTGGTGTGAGAGGATGATGATGATGATGATGATG 92673  
 301 GCCGAGCTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 355  
 92672 GCCGAGCTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 92613  
 356 TGATCCAGCAATGCGCGGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 415  
 92612 TGATCCAGCAATGCGCGGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 92553  
 416 GAAAGAAAGGTTACGCTAATTAATGTCGATCCCATGACGCTTCCGATGCTTCACTG 475  
 92552 GAGGAAAGGTTACGCTAATTAATGTCGATCCCATGACGCTTCCGATGCTTCACTG 92493  
 476 CTAATCTAGTCCGACGCGCGGCTTAATGCTAGGAGGAGGAGGAGGAGGAGGAGG 535  
 92492 CTAATCTAGTCCGACGCGCGGCTTAATGCTAGGAGGAGGAGGAGGAGGAGGAGG 92433  
 536 GCGCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 595  
 92432 GCGCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 92373  
 596 GGAATGCGTTGAACTCAAGCTAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGG 655  
 92372 GGAATGCGTTGAACTCAAGCTAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGG 92313  
 656 AGCAGTGAATGCGTGAAGATTAATGAGAAATGATGAGGAGGAGGAGGAGGAGGAGG 715  
 92312 AGCAGTGAATGCGTGAAGATTAATGAGAAATGATGAGGAGGAGGAGGAGGAGGAGG 92253  
 716 AACTGAGCTCATGACGAAAGCTGGGAGCAACAGATTAGATCCCTGCTGATGCTC 775  
 92252 AACTGAGCTCATGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 92193  
 776 AGCGCTTAACGATGCACTAGTTGTTGGGCTTAT-TAGGCTTGTGAAGAGCTTAAC 834  
 92192 AGCGCTTAACGATGCTTACAGTGTGGGCTTATTAAGAGCTTAAGAGGAGGAGGAGG 92133

835 GCGTGAAGTTGACCGCTGGGAGGATACGCTGACAGATTAAACTCAAGATTGACGG 894  
 92132 GCAATTAAGTGAACCGCTGGGAGGATACGCTGACAGATTAAACTCAAGATTGACGG 92073  
 895 GGACCGGACCAAGCGGATGATTAATGATTAATGATGATGATGATGATGATGATGAT 954  
 92072 GGACCGGACCAAGCGGATGATTAATGATTAATGATGATGATGATGATGATGATGAT 92013  
 955 ACCCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1013  
 92012 GGTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91953  
 1014 GGTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073  
 91952 GGTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91893  
 1074 GGTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1132  
 91892 GGTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91833  
 1133 ACAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1192  
 91832 ACAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 91773  
 1193 ACAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1252  
 91772 ACAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 91713  
 1253 AAAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312  
 91712 AAAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91653  
 1313 GATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1372  
 91652 GATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91593  
 1373 GATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1432  
 91592 GATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91535  
 1433 GATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
 91534 GATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91510

RESULT 51  
 US-09-557-884-1  
 Sequence 1, Application US/09557884  
 Patent No. 6506581  
 GENERAL INFORMATION:  
 APPLICANT: Fleischmann et al.  
 TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof  
 NUMBER OF SEQUENCES: 1  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: MD  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3 1/2 inch diskette  
 COMPUTER: Dell Pentium  
 OPERATING SYSTEM: MS DOS V6.22  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/557,884  
 FILING DATE: 25-Apr-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:

```
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marka
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match      66.7%; Score 971.8; DB 3; Length 1830121;
Best Local Similarity 81.3%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 267; Indels 7; Gaps 5;

QY      1 ATTGAACGCTGGCGGAGCTGTTTACATGACATGCAAGTGAACGCGACGACG--ATGCTTGC 58
DB      770639 ATTGAACGCTGGCGGAGCTTAAACATGCAAGTGAACGCGTAAAGGAAAGCTTGC 770698
QY      59 ATC--TGATGCGAGATGGCGGAGCGGAGTGAATGCAATCGAAACGTATCCAAAGAGGGG 116
DB      770699 TTTCTTGCTGACAGATGGCGGAGCGGAGTGAATGCTTGGGAATCGGCTTATGAGGGG 770758
QY      117 GGTAAACGATCGAAAGATGCTAATATACGATATCTCTAAAGAGAAAGACGGGGATC 176
DB      770759 GATTAACGAGGAAACCTGTCCTAATACGCGCTAATATGCGAAGATGAAGATGCGGAGCT 770818
QY      177 GAAAGACCTTGGCGCTTTTGAAGCGGCGGATGCTGAATGACTAATGATTGGTGAAGGC 236
DB      770819 GAGAGCGCGCATGCCATAGGATGAGCCAAAGTGGATTAAGTATGGTGGGTTAAATGC 770878
QY      237 CTACCAAGCGGACGATCAGTATGTTGCTGAGAGAGACACAGCCCACTGGAGCTGAAGA 296
DB      770879 CTACCAAGCGCTGATCTCTAGCTGCTGAGAGATGACACAGCCCACTGGAACTGAAGA 770938
QY      297 CACGCGCCAGACTCCTACGCGGAGGAGCAGAGTGGGGAAATTTTGGCAATGGGGCGACAGCT 356
DB      770939 CACGCTCCAGACTCTTACGCGAGGAGCAGAGTGGGGAAATTTGGCAATGGGGCGACAGCT 770998
QY      357 GATCCAGCAATGCGCGGTAGTGAAGAGAGCCCTTCGCGTGTAAAGCTCTTTCAGTCCAG 416
DB      770999 GACGCGCATGCGCGGTGAATGAAGAGCCCTTCGCGTGTAAAGCTCTTTCGGTATTCG 771058
QY      417 AAGAAAGGTTACGGTAAATATCTGACCCATGACGCTATGACAGAGAGACAGACCGGC 476
DB      771059 ACGAAGGTTGATGTGTTAATAGACATCAAAATTTGAGCTTAAATACAGAAAGACACCGGC 771118
QY      477 TAACATACGTCGACGACGCGCGGTATATAGTAGGGGTCGCAAGGCTTAAATGGGAATTAC 536
DB      771119 TAACCTCGTCGACGACGCGCGGTATATAGAGAGGTCGCAAGGCTTAAATGGGAATTAC 771178
QY      537 GCGTAAAGGAGTGCAGAGCGCGCTTGTAAAGTCAGATGTAAATCCCGCGGCTTAAACCTCG 596
DB      771179 GCGTAAAGGAGTGCAGAGCGCGGTATTTAATGAGGTGTAAGAGCCCTCGGCTTAACTC 771238
QY      597 GAATTGCGTTTGAACACTACAAAGCTAGATGTGGCAGAGGAGGTGAATTTCCATGTGTA 656
DB      771239 GAATTGCAATTCAGACTGGGTAACTAAGATACTTTAGGAGGAGGTGAATTTCCACGTGTA 771298
QY      657 GGAAGTAAATGGGTAGATATGAAAGACATCGATGGGAGGAGCGCTCCGCGGTAA 716
DB      771299 GCGGTGAATATGCGTAGATATGAGAGAAATACGAAAGCGAGCGCCCTTGGGAAATG 771358
QY      717 CACTGACGCTCATGACGAAAGCGTGGGAGCAACAGAGATTAGATACCTGTTAGTCA 776
DB      771359 TACTGACGCTCATGTCGGAAGCGTGGGAGCAACAGAGATTAGATACCTGTTAGTCA 771418
```

```
QY      777 CGCCCTAAACGATGCACTAGTGTGGGCGCTTATTAGGCTTGGTAACGAAGCTAACGC 836
DB      771419 CCTGTAAACGCTGTGCTGATTTGGGATTTGGGCTTAA-GAGCTTGTGCGCTGATACCT 771477
QY      837 GTGAAGTTGACCGCGCTGGGAGTACGATCGACAAATTAATACTCAAGAAATTTGACGGGG 896
DB      771478 GATTAATGACCGCGCTGGGAGTACGCGCGCAAGTTAAATCTCAATGAATTTGACGGGG 771537
QY      897 ACCCGACAAAGCGGTGATTAATGAGATTAATTCATGCAACGCGAATAAACCCTTACTAC 956
DB      771538 GCCCGACAAAGCGGTGAGCATGTGTTAATTCATGCAACGCGAATAAACCCTTACTAC 771597
QY      957 CCTTGACATGTAAGCCAAATTTTCTAGAGAT-AATTAAGTCTTCGGGAACGCTTAACACAG 1015
DB      771598 TCTTGACATCTTAAGAAAGACTCAAGAGTGAACCTGTGCTTCGGGAACCTTAAGAGCAG 771657
QY      1016 TGTGCAATGCTGTGCTGACAGCTGTCGTCGATGAGATGTTGGGTTAAGTCCCGACAGCG 1075
DB      771658 TGTGCAATGCTGTGCTGACAGCTGTCGTCGATGAGATGTTGGGTTAAGTCCCGACAGCG 771717
QY      1076 CAACCTTGTCTAATTAATTCATC-ATTTGGTTGGCACTTTAATGAGACTGCCGTGAC 1134
DB      771718 CAACCTTAATCTTTGTTGCGACGACCTTGTGCGGAACTCAAGAGACTGCCAGTGAT 771777
QY      1135 AAACCGGAGAGTGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGCGTTAC 1194
DB      771778 AAACCTGAGAGAGTGGGGATGACGTCAAGTATATATGAGCCCTTAAGAGAGGCTACAC 771837
QY      1195 ACGTAATACATGAGCGCGTACAGAGGGTTGCCAACCCGAGGGGAGCTAATCTCAGAA 1254
DB      771838 ACGTCTACAAATGGGCTATACAGAGGAAAGCAACCTCGAGTGGAGGAATCTCATAA 771897
QY      1255 AGCGGTGTGATCCGATCGAGATCGAATCGAATCGAATCGAATCGAATCGCTAGT 1314
DB      771898 AGTACGTCAAGTCCGAGATGGAATGCTCACTCACTCAATGAAGTGGAAATCCCTAGT 771957
QY      1315 AATCGCGATACGATGTCGCGGTGAATACGTTCCCGGGCTCTTGAACACACCGCCGCTCA 1374
DB      771958 AATCGCGATACGATGTCGCGGTGAATACGTTCCCGGGCTCTTGAACACACCGCCGCTCA 772017
QY      1375 CACCATGAGAGTGGGTTTACACGAAAGCAGGATCTAACCGTAAAGAGAGCGCTTGCCA 1434
DB      772018 CACCATGAGAGTGGGTTTACACGAAAGTAAATGATTAATCTTTTGGAGGGGCTTACCA 772077
QY      1435 CGGTGATTCATGACTGCGGGTG 1457
DB      772078 CGGTATGATTCATGACTGCGGGTG 772100

RESULT 52
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fieischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome. Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US/09/557,884
/ FILING DATE: 25-Apr-2000
/ CLASSIFICATION: <Unknown>
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: 08/476,102
/ FILING DATE: JUN-5-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Michelle S. Marks
/ REGISTRATION NUMBER: 41,971
/ REFERENCE/DOCKET NUMBER: PB186P3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 301-309-8504
/ TELEFAX: 301-309-8439
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1830121 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match      66.7%; Score 971.8; DB 3; Length 1830121;
Best Local Similarity 81.3%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 267; Indels 7; Gaps 5;

QY 1 ATTGAAGCTGCGCGCATGCTTTACACATGCAAGTCGAAACGCGACGACGCG-ATGCTTGC 58
DB 128095 ATTGAACGCTGCGCGCATGCTTTACACATGCAAGTCGAAACGCGACGCG-ATGCTTGC 128036
QY 59 ATC--TGTGCGCGAGTGGCGGACGCGGTGAGTATGTCGACGACGTAATCCAGAAAGAGGG 116
DB 128035 TTTCTTGCTGACGAGTGGCGGACGCGGTGAGTATGTCGACGACGTAATCCAGAAAGAGGG 127976
QY 117 GGTAAACGATCGAAAGATGCTTAAATCCGATATCTTAAGGAGGAAAGAGCGGGATC 176
DB 127975 GATTAACGACGCGGAAACTGTGCTTAATACCGGTAATATCGGAGATGAAATGCGGAGCT 127916
QY 177 GAAAGACCTTGCGCTTTTGGAGCGCGCATGCTGATTAAGCTAGTTGGTGGGTAAGGC 236
DB 127915 GAGAGCGCGCATGCTTAAAGATGAGCCCAAGTGGGATTAAGTATTTGGTGGGTAAGGC 127856
QY 237 CTACCAAGCGACGATCAGTATGTTGTTCTGAGAGACGACGACCACTGGGACTGAGA 296
DB 127855 CTACCAAGCGTGCATCTCTAGTGTCTGAGAGATGACGACGACCACTGGGACTGAGA 127796
QY 297 CACGCGCCAGACTCTTACGCGGAGCGACGATGCGGAAATTTTGGACATATGGCGCAAGCT 356
DB 127795 CACGCTCAGACTCTTACGCGGAGCGACGATGCGGAAATTTTGGCGCAATGCGGAGCAAGCT 127736
QY 357 GATCACAAGATGCGCGCGTGAAGAGAGCGCTTGGGTTGAAAGCTCTTCAAGTCGAG 416
DB 127735 GACGACGATGCGCGCGTGAAGAGAGAGCGCTTGGGTTGAAAGCTCTTCAAGTCGAG 127676
QY 417 AAGAAAAGGTTACCGTAAATATATGTCGACCCATGACGCGTATCGACAGAAAGACCGGC 476
DB 127675 AAGAAAGGTTATGTTAATATGTCATCAAAATTTGACGTTAATACAGAAAGACCGGC 127616
QY 477 TAACTACGTGCGACGACCGCGGTATACGTAAGGTTGCAACGTTAATCCGAATTAATCGG 536
DB 127615 TAACTCCGTGCGACGACCGCGGTATACGTAAGGTTGCAACGTTAATCCGAATTAATCGG 127556
QY 537 GCGTAAAGGCGCGACGCGCGCTTGAATGTCAGATGTAATCCCGGCGCTTAACTGAG 596
DB 127555 GCGTAAAGGCGCGACGCGCGCTTGAATGTCAGATGTAATCCCGGCGCTTAACTGAG 127496
QY 597 GAATTGCGTTTGAACCTAACAGCTAGAGTGTGCGACGAGGAGGTGAATTCATGTGTA 656
DB 127495 GAATTGCAATTTCAAGCTGGTAACTAGAGTACTTTAAGGAGGAGGTGAATTCATGTGTA 127436
QY 657 GCAGTGAATGCGTGAATATGAGAAACATCATGCGCAAGGCGACGCTCTGGGTAA 716
DB 127435 GCGGTGAATGCGTGAATATGAGAAACATCATGCGCAAGGCGACGCTCTGGGTAA 127376
```

```
QY 717 CACTGACGCTATGACGAGAAAGCGTGGGAGCAAAAGATTAAGTATCCCTGTAAGTCCA 776
DB 127375 TACTGACGCTATGTCGCAAAAGCGTGGGAGCAAAAGATTAAGTATCCCTGTAAGTCCA 127316
QY 777 CCGCCCTAAACGATGTCACCTAGTGTGCGCCTTATTAAGCCTTGTGAACGACCTAACGC 836
DB 127315 CCGCTAAACGCTGTGCTGATTTGGGAGTTGGGCTTA-GAGCTGTGTCCCGTGTACCTG 127257
QY 837 GTGAAGTTGACCGCCTGGGAGTACCGTGGCAAGTTAAACCTCAAGAAATGAAGAGGGG 896
DB 127256 GATTAATCGACCGCCTGGGAGTACCGCAGGTTAAACCTCAATGATGATGAGCGGG 127197
QY 897 ACCCGCAAGCGGTGATTAATGTTGATTAATTCGATGACGCAAGCAAAACCTTACTAC 956
DB 127196 GCCCGCAAGCGGTGAGATGTTAATTCGATGACGCAAGCAAAACCTTACTAC 127137
QY 957 CTTGACATGATGACGAATTTTCTAGAGAT-AGATTAGTGTTCGGGAACGCTAACACAGG 1015
DB 127136 TCTTGACATCTTAAGAAAGCTCAGAGATGAGCTTGTGCTTGGGAACTTAAGACAGG 127077
QY 1016 TGTGATGAGCTGTGCTGACCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGACG 1075
DB 127076 TGTGATGAGCTGTGCTGACCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGACG 127017
QY 1076 CAACCTTGTCTAATATGTCATC-ATTGTTGGGCACTTATGAGACTGCGCGTGAC 1134
DB 127016 CAACCTTATCTTGTGTGTCAGACGACTGTGTGCGGAATCAAAAGAGACTGCAAGTAT 126957
QY 1135 AAACCGGAGAGGTGGGAGTACGTCAGTCTCATATGCGCTTATGAGTGGGCTTCAAC 1194
DB 126956 AAACCTGAGAGAGGTGGGAGTACGTCAGTCTCATATGCGCTTATGAGTGGGCTTCAAC 126897
QY 1195 ACCTAATACATGCGCGGTACAGAGGTTGCCAACCGCGAGGGGAGCTTAATCTCAGAA 1254
DB 126896 ACCTGATCAATGCGCGGTACAGAGGAGCGAAGCTGCGAGGTGAGCGAATCTCAATA 126837
QY 1255 AGCGCGCTGTAATGCGGATGCGAGTGTGCAACTGCACTCCGTGAAGTCGGAATGCTAGT 1314
DB 126836 AGTACGCTTAATGCTGCGGATGAGTGTGCAACTGCACTCCGTGAAGTCGGAATGCTAGT 126777
QY 1315 AATCGCGATCAGATGTCGCGGTGAATATAGTTCCCGGCTTGTATACACCGCGCTCA 1374
DB 126776 AATCGCAATCAGATGTCGCGGTGAATATAGTTCCCGGCTTGTATACACCGCGCTCA 126717
QY 1375 CACCATGAGAGTGGGTTTACAGAAACAGTAGTCTTAACCTGAAGAGGCGCTTCCCA 1434
DB 126716 CACCATGAGAGTGGGTTTACAGAAAGTATAGCTTAACCTTTGGAGGGCGCTTACCA 126657
QY 1435 CGGTGAGATTCAATGCTGGGGTG 1457
DB 126656 CGGTATGATTCAATGCTGGGGTG 126634
```

```
RESULT 53
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
```



```

;
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: FBI66P1C1
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match      66.7%; Score 971.8; DB 3; Length 1830121;
Best Local Similarity 81.3%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 267; Indels 7; Gaps 5;

QY 1 ATTGAACGTGGGGGATGCTTTACATGCAAGTGGAAACGGGACGCG--ATGCTTGC 58
DB 770639 ATTGAACGTGGGGGATGCTTTACATGCAAGTGGAAACGGGACGCGAGTGG 770698
QY 59 ATC--TGTGTGCGAGTGGCGGACGCGGTGAGTAATGCAATCGAAGAGGGG 116
DB 770699 TTTCTTGTGACGAGTGGCGGACGCGGTGAGTAATGCTTGGGAATCTGGCTTAATGAGGGG 770758
QY 117 GGTAAACGATCGAAAGATGCTTAATACCGCATATATCTTAAGAGAGAAACAGCGGGATC 176
DB 770759 GATTAACGAGCGGAAACTGTCTTAATACCGCATATATCGGAAAGATGAAAGATGCGGAGCT 770818
QY 177 GAAAGACCTTGGGCTTTTGGAGCGCGCATGCTGATTAAGCTAGTTGGTGGGGTAAAGGC 236
DB 770819 GAGAGGCGCCATGCTCAATAGGATAGCCCAAGTGGGATGAGTATGTTGGTGGGTTAAATGC 770878
QY 237 CTAACCAAGGCGACGATCAGTATGTTGCTGAGAGAGACGACCAACCACTGGGACTGAGA 296
DB 770879 CTAACCAAGGCTGAGATCTCTAGTGTGTTGAGAGAGTGAACGACCACTGGAACAGAGA 770938
QY 297 CACGCGCCAGACTCTTAACGAGGAGGACAGAGTGGGAAATTTTGGACAATGGCGCAACCT 356
DB 770939 CACGCGCCAGACTCTTAACGAGGAGGACAGAGTGGGAAATTTTGGCGCAATGGCGGAAACCT 770998
QY 357 GATCCAGCAATCCCGGTGAGTGAAGAAGGCTTGGGTTGTAAACCTCTTTCAGTCGAG 416
DB 770999 GACGCAAGCATGCGCGGTGAATGAAGAAGGCTTGGGTTGTAAATCTTTCGGAATTG 771058
QY 417 AAGAAAGGTTACGGAATTAATCGTACCACGATGAGTATCGACAGAAAGACCGGC 476
DB 771059 AAGAAAGGTTGATGTGTTAATAGTACATCAATTTGACGTTAAATCAAGAAAGACCGGC 771118
QY 477 TAACTAAGTGCAGACGCGCGGTAAATACGTAGGAGTGCAGACGCTTAATCGAAATTAAC 536
DB 771119 TAACTCGTCGACGACGCGCGGTAAATACGTAGGAGTGCAGACGCTTAATCGAAATTAAC 771178
QY 537 GCGTAAGAGGTGCGCAGCGGCGCTTGTAACTGATGTAAATCCCCGGGCTTAACCTGG 596
```

```

DB 771179 GCGTAAGAGGACGACGAGCGGCTTATTTAAGTAGAGTGTGAAGCCCTGGGCTTAACCTAG 771238
QY 597 GAATTGCGTTTGAACCTACAAAGCTAGATGTGGCAGAGGAGGTGGAATTTCCATGTGTA 656
DB 771239 GAATTGCAATTCAGACTGGGTCACTAGAGTACTTTAGGAGGGGTGAATTTCCAGCTGTA 771298
QY 657 GCAGTGAATGCGGTAGATATGAAAGAACATGATGGGAGGACGCTCTGGGTTAA 716
DB 771299 GCGGTGAATGCGGTAGATATGAGAGATACCGAAGGCGAAGGACGCCCTTGGGAATG 771358
QY 717 CACTGACGCTCATGACAGCAAGACGTTGGGAGCAACAGATTAGTACCCTGTAAGTCA 776
DB 771359 TACTGACGCTCATGAGCGCAAGACGTTGGGAGCAACAGATTAGTACCCTGTAAGTCA 771418
QY 777 CGCCCTAAACGATGTCAACTAGTTGTTGGGCTTTATGAGCTTGGTATACGAAGCTAAC 836
DB 771419 CGCTGTAAACGCTGTGATTTGGGATTTGGGCTTTA-GAGCTTGTGCGCCGTGCTAAACGT 771477
QY 837 GTGAAGTTGACCGGCTGGGAGTACGTTGCGAAGTTAAACTCAAGGAATTGACGGGG 896
DB 771478 GATTAATGACCGGCTGGGAGTACGCGCGCAAGGTTAAACTCAAGGAATTGACGGGG 771537
QY 897 ACCGCAACGCGGTGATTTATGATGATTAATTCATGCAACGCGAAGAAACCTTACTAC 956
DB 771538 GCCCGCAACGCGGTGAGCATGTGCTTTAATTCATGCAACGCGAAGAACTTACTAC 771597
QY 957 CTTGACATGATGCAATTTTCTAGAT-AGATTAGTCTTGGGAAACCTTAACACAG 1015
DB 771598 TCTTGAACATCTTAAGAAAGCTCAAGATGAGCTTGTGCTTGGGAACTTAAGACAG 771657
QY 1016 TGTGCAATGCGGTGTGCTGACGCTGTGTGATGATGTTGGGTTAATGCCGCAAGAG 1075
DB 771658 TGTGCAATGCGGTGTGCTGACGCTGTGTGATGATGTTGGGTTAATGCCGCAAGAG 771717
QY 1076 CAACCTTGTCAATTAATTCGCATC-ATTTGTTGGGCACTTTAATGAGACTGCGGTGAC 1134
DB 771718 CAACCTTAATCTTTGTTGTCGACGACCTTGTGTGGGAACTCAAGAGACTGCGGTGAT 771777
QY 1135 AAACCGAGGAAGTGGGGATGATGCTCAAGTCTCATGAGCTTATGGTATGAGCTTAC 1194
DB 771778 AAACCTGAGGAAGTGGGGATGATGCTCAAGTCTCATGAGCTTATGGTATGAGCTTAC 771837
QY 1195 AGCTAATCAATGAGGCGGTACAGAGGGTTGCCAACCGCGAGGGGAGCTAATCTCAGAA 1254
DB 771838 AGCTAATCAATGAGGCGGTACAGAGGGTACAGAGGGTACAGAGGGTACAGAGGGTAC 771897
QY 1255 AGCGGCTGTAGTCCGATTCGAGTCTGCACTCGACTCCGTAAGTCCGTAAGTCCGTA 1314
DB 771898 AGTACGTCAATGCTCCGATTCGAGTCTGCACTCGACTCCGTAAGTCCGTAAGTCCGTA 771957
QY 1315 AATCGCGATCAAGATGTCGCGGTAAATCGTTCCCGGGCTTGTAAACACCGCCGCTCA 1374
DB 771958 AATCGCGATCAAGATGTCGCGGTAAATCGTTCCCGGGCTTGTAAACACCGCCGCTCA 772017
QY 1375 CACCAATGGAAGTGGGTTTCAACGAAGCAAGTAACTAACCGTAAAGGAGGCGTTGCCA 1434
DB 772018 CACCAATGGAAGTGGGTTTCAACGAAGCAAGTAACTAACCGTAAAGGAGGCGTTGCCA 772077
QY 1435 CGGTGAGATTCACTGAGGAGTG 1457
DB 772078 CGGTGAGATTCACTGAGGAGTG 772100

RESULT 54
US-09-643-990A-1/c
; Sequence 1, Application us/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
;
```



J. Craig Venter  
TITLE OF INVENTION: The Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
thereof, and Uses Thereof  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville,  
STATE: MD  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/643,990A  
FILING DATE: 23-Aug-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,429  
FILING DATE: 1995-06-07  
APPLICATION NUMBER: 08/426,787  
FILING DATE: 1995-04-21  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB186P1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-610-5790  
TELEFAX: 310-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-643-990A-1  
Query Match 66.7%; Score 971.8; DB 3; Length 1830121;  
Best Local Similarity 81.3%; Pred. No. 0;  
Matches 1189; Conservative 0; Mismatches 267; Indels 7; Gaps 5;  
QY 1 ATGGAAGCTGGCGGCAATGCTTTACATGCAAGTCCGAACGGCAGCAGCG--ATGCTTGC 58  
DB 128095 ATGGAAGCTGGCGGCGGCTTTACATGCAAGTCCGAACGGTACGAGGAAAGCTTGC 128036  
QY 59 ATC--TGGTGGCGAGTGGCGGAGCGGTGAGTAATGCAATCGGAAGCTATCCAGAAAGAGGG 116  
DB 128035 TTCTCTTCTACGAGTGGCGGAGCGGTGAGTAATGCTTGGGAATCTGGCTTATGAGAGGG 127976  
QY 117 GGTAAACGATCGAAAGATGTGCTAATACCGCATCTACTTAAGAGGAAAGCAGGGGATC 176  
DB 127975 GATTAACGAGCGGAAATCTGTGCTAATACCGGATTTATTCGGAATGAAAGTGGGAGCT 127916  
QY 177 GAAAGACCTTGGCGCTTTTGAAGCGCGGAGTCTGATTAAGTCTAGTTGGTGGGTTAAAGGC 236  
DB 127915 GAGAGCGCGCATGCTATGATGATGAGCCCAAGTGGATTAAGTGTGGTGGGTTAAATGC 127856  
QY 237 CTACCAAGCGGAGCATGATGATGTTGTTGCTGAGAGGAGCAGCAGCAGCAGCTGGAGCTGAG 296  
DB 127855 CTACCAAGCGTGCATCTCTAGCTGTGTGAGAGATGACCAAGCAGCAGCTGGAATCTGAG 127796  
QY 297 CACGCGCCAGACTCTCTACGAGGAGCAGCATGTGGGAAATTTTGAACAATGGCGCAGAGCT 356  
DB 127795 CACGCTCAAGACTCTCTACGAGGAGCAGCATGTGGGAAATTTTGCAGATGGGGGAAACCT 127736  
QY 357 GATCAGCAATGCGCGGTGAGTGAAGAGGCTTCCGCTTGTAAAGCTCTTTCAAGTGAAG 416  
DB 127735 GACGAGCGCATGCGCGGTGATGAAGAGGCTTCCGCTTGTAAAGTCTTTTCAGTATTCG 127676

QY 417 AAGAAAAGTTACGTTAAATTAATCGGACCAGCCGATATGACAGAAAGACACCGGC 476  
DB 127675 AAGAAAGTTGATGTGTAAATAGTACATCAATATGAGTTAAATACGAAAGAGACCGGC 127616  
QY 477 TAACTAGTGCAGCAGCGCGGTAATACGTAGGTTGCAAGCGTTATTCGAAATTAATCTGG 536  
DB 127615 TAACTCGGTCCAGCAGCGCGGTAATACGTAGGTTGCAAGCGTTATTCGAAATTAATCTGG 127556  
QY 537 GCGTAAAGGAGTGGCAGCGCGCTTGTAAAGTCAAGATGTGAATTCCTCGGCTTAACCTGG 596  
DB 127555 GCGTAAAGGAGCAGCAGCGGTTATTAAGTGGTGTGAAGCCCTGGGCTTAACCTGG 127496  
QY 597 GAATTCGCTTTGAATACTACAAAGCTGAGTGTGCGCAGAGGAGGTGAATTCGATGTGA 656  
DB 127495 GAATTCATTTTCAGACTGGGTAACTAGAGTACTTTTAGGAGGGGTGAATTCGATGTGA 127436  
QY 657 GCAGTAAATGCGTAGATATATGAGAAATATGATGAGGAGGAGCGCTCTGGGTTAA 716  
DB 127435 GCGGTAAATGCGTAGATATATGAGAAATATGAGAGAAATCCGAGGCGGAGCGAGCTTGGGAATG 127376  
QY 717 CACTGACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAATACCTGTGATGCCA 776  
DB 127375 TACTGACGCTCATGTCGAAAGCGTGGGAGCAAAACAGATTAATACCTGTGATGCCA 127316  
QY 777 CGCCTTAACGATGTCACTAGTTGTGGGCTTTATTAAGCTTGTGAACGAACTAACCG 836  
DB 127315 CGCTGTAACGCTGTGATTTGGGAGTGGGCTTTA--GAGTGTGTGCTCCGTACTTAACGT 127257  
QY 837 GTAAAGTTGACCGCGTGGGAGTACGAGTGCAGATTTAAATCAAGAAATGACGAGG 896  
DB 127256 GATTAATTCAGCCGCTGGGAGTACGAGCGGAGGTTAAATCAAGAAATGACGAGG 127197  
QY 897 ACCCGCAAGCGGTGATTAATGATTAATTCGATGCAACGCGAAATCTTAACCTAC 956  
DB 127196 GCGCGCAAGCGGTGAGATGTGTTAATTCGATGCAACGCGAAATCTTAACCTAC 127137  
QY 957 CTTGACATGTAGCAATTTTCTAGAGAT--AGTTAGTGTCTTGGGAAACGCTAACACAG 1015  
DB 127136 TCTTGACATCTCTAAGAAAGCTCAGAGATGAGTGTGCTTGGGAACTTAAGACAGG 127077  
QY 1016 TGCTGATGCGTGTGCTGAGCTGTGTGCGTGAATGATGGGTTAAGTCCGCAACGAGC 1075  
DB 127076 TGCTGATGCGTGTGCTGAGCTGTGTGCGTGAATGATGGGTTAAGTCCGCAACGAGC 127017  
QY 1076 CAACCTTGTCTAATTAATGCGATC--ATTGCTTGGGCACTTTATGAGACTGCGCGTAC 1134  
DB 127016 CAACCTTATCTTTTGTGCAACGACTTGTGCGGAACTCAAGAGACTGCGCAATGAT 126957  
QY 1135 AAACCGGAGAGGTGGGATGACGTCAAGTCTCATGCTCTTAATGAGGCTTAC 1194  
DB 126956 AAACCTGAGAGGTGGGATGACGTCAAGTCTCATGCTCTTAATGAGGCTTAC 126897  
QY 1195 ACGTAATACATGCGCGGTACAGAGGTTGCCAACCCTGGAGGGAGCTAATCTCGAA 1254  
DB 126896 ACGTCTACATGCGCGGTACAGAGGTTGCCAAGCTGCGAGGTTGAGCGCAATCTCAA 126837  
QY 1255 AGCGCGTGTAGTCCGATCGGATCGAAGTCTGCAACTGCACTCCGTGAAGTCCGAAATCGTACGT 1314  
DB 126836 AGTACGTTCTAATGTCGAGTGTGAGTCTGCAACTGCACTCATGAAGTCCGAATGCTACGT 126777  
QY 1315 AATCGCGATCAGCATGTGCGGTTGAATACGTTCCCGGCTTTGTACACACCGCCGTCA 1374  
DB 126776 AATCGCGATCAGCATGTGCGGTTGAATACGTTCCCGGCTTTGTACACACCGCCGTCA 126717  
QY 1375 CACCATGGAGTGGCTTTACAGAAAGCTGTGTCTAACCCTGAAGAGGCGCTTGCCA 1434  
DB 126716 CACCATGGAGTGGCTTTACAGAAAGCTGTGTCTAACCCTTTTGAAGGCGCTTTACCA 126657  
QY 1435 CCGTGAATTCATGACTGGGGT 1457  
DB 126656 CCGTGAATTCATGACTGGGGT 126634

```
RESULT 55
US-10-158-865-1
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Pleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2CID1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40810)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55309)..(55309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55359)..(55359)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
```

```
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
Query Match 66.7%; Score 971.8; DB 3; Length 1830121;
Best Local Similarity 81.3%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 267; Indels 7; Gaps 5;
1 ATTGAACCTGCGCGGCACTTTACATGCAAGTGCAGACGAGCG--ATGCTTGC 58
770639 ATTGAACCTGCGCGGCACTTTACATGCAAGTGCAGACGAGCGAGAACTTGC 770698
59 ATC--TGATGCGAGTGGCGGCGAGTGAATGATGATGCAAGCTATCCAGAGAGGGG 116
770659 TTCTTGTGCTGACGATGGCGGAGCGGCTGATGCTTGGGAACTGGCTTATGAGGGG 770758
117 GGTAAAGCATCGAAAGATGTGCTAATACCGCATATCTTTAAAGAGAAAGCGGATC 176
770759 GATTAAGACGGGAACTGTGCTAATACCGCATATCTTAAAGAGATGAAGTGGCGGACT 770818
```

```
177 GAAAGACCTTGCGCTTTTGGAGCGGCGATGCTGATTAGTAGTGTGGGTAAAGGC 236
770819 GAGAGCGCGCATGCGCATGAGATGAGCCCAAGTGGATTAAGTAGTGGGTAAATGC 770878
237 CTACCAAGCGCAGATCAGTAGTGTGCTGAGAGACGACCACTGCGACTGAGA 296
770879 CTACCAAGCTGCGATCTCTAGCTGTCTGAGAGATGACGACCACTGGAACCTGAGA 770938
297 CACGCGCCAGACTCTTACCGGAGGACAGCTGGGGAATTTTGAACATGGGCGCAAGCT 356
770939 CACGCTCAGACTCTTACCGGAGGACAGCTGGGGAATTTTGAACATGGGCGCAAGCT 770998
357 GATCCAGCAATGCGGTGATGAGGAAGGCGCTGGGTGTAAGCTCTTCAGTCCAG 416
770999 GACGACGCGATCGCGGTGATGAGGAAGGCGCTGGGTGTAAGCTCTTCAGTCCAG 771058
417 AAGAAAGGTTACGGTAATTAATCGTGAACCCATGACGATGACAGAAAGACCGGC 476
771059 AAGAAAGGTTACGGTAATTAATCGTGAACCCATGACGATGACAGAAAGACCGGC 771118
477 TAACTACGTGCGACGACCGCGGTAAATACGTAGGCTGCAAGCTTAATGGAATTA 536
771119 TAACTACGTGCGACGACCGCGGTAAATACGTAGGCTGCAAGCTTAATGGAATTA 771178
537 GCGTAAAGGCTGCGGAGCGGCGCTTGTAGTCAAGTGAATCCCGGGCTTAACCTG 596
771179 GCGTAAAGGCTGCGGAGCGGCGCTTGTAGTGAAGTGAAGCTTGGGCTTAACCTG 771238
597 GAATTCGCTTGAACCTAACAAGCTAGAGTGTGCGAGAGGTGGAATTCATGTGTA 656
771239 GAATTCGCTTGAACCTAACAAGCTAGAGTGTGCGAGAGGTGGAATTCATGTGTA 771298
657 GCAGTGAATGCGTAGAGATGAGAAACATCATGCGGACGACCTCTCGGTTAA 716
771299 GCAGTGAATGCGTAGAGATGAGAAACATCATGCGGACGACCTCTCGGTTAA 771358
717 CACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTGATACCTGATGCA 776
771359 TACTGACGCTCATGAGTGAAGGTGGGAGCAACAGATTGATACCTGATGCA 771418
777 CGCCCTTAAAGATGTAACATGATGTTGGCTTATTAAGCTGTGTAACGATGCA 836
771419 CGCCCTTAAAGATGTAACATGATGTTGGCTTATTAAGCTGTGTAACGATGCA 771477
837 GTGAAGTTGACCGCTGCGGAGTACGCTCGCAAGTTAAATCTCAAGAAATGACGGG 896
771478 GTGAAGTTGACCGCTGCGGAGTACGCTCGCAAGTTAAATCTCAAGAAATGACGGG 771537
897 ACCCGCAAGCGGTGATTAATGATGATTAATTCGATGCAACGCAAAACCTTAC 956
771538 ACCCGCAAGCGGTGATTAATGATGATTAATTCGATGCAACGCAAAACCTTAC 771597
957 CTTGACATGATGGAATTTTCTAGAGAT--AGATTAGTGTGCGGAACGTAACAGG 1015
771598 CTTGACATGATGGAATTTTCTAGAGAT--AGATTAGTGTGCGGAACGTAACAGG 771657
1016 TGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075
771658 TGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771717
1076 CAACCTTGTATTAATTCGATC--ATTGCTTGGGCACTTTAATGAGTCCGCGTAC 1134
771718 CAACCTTGTATTAATTCGATC--ATTGCTTGGGCACTTTAATGAGTCCGCGTAC 771777
1135 AAACCGAAGAGGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1194
771778 AAACCGAAGAGGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 771837
1195 ACGTAAATGACGCGTACAGAGGTTGCAACCGCGAGGGGAGCTTAATTCAGAA 1254
771838 ACGTAAATGACGCGTACAGAGGTTGCAACCGCGAGGGGAGCTTAATTCAGAA 771897
```

QY	1255	GGCGGTGTAAGTCCGGATCCGAGTCTGCACTCGAATCCGGAGTCCGTAAGT	1314
Db	771988	AGTACGTCTAAGTCCGGATGGAGTGTGCACTTCGATGAAGTGGAAATCGTAACT	771957
QY	1315	AATCCGGATCGCATGTGTGCGGTGAATACGTTCCCGGGTCTTGACACCGCCGTC	1374
Db	771958	AATCGCAATCAAGATGTGTGCGGTGAATACGTTCCCGGGTCTTGACACCGCCGTC	772017
QY	1375	CACCATGGAGTGGGTTTACCAAGACAGTAACTGTAACCGTAAGAGGGCGCTTGCCA	1434
Db	772018	CACCATGGAGTGGGTTTACCAAGACAGTAACTGTAACCGTTTGGAGGGCGTTACCA	772077
QY	1435	CGGTGAGATTCATGACTGGGGGTG	1457
Db	772078	CGTATGATTCATGACTGGGGGTG	772100
RESULT 56			
US-10-158-865-1/c			
: Sequence 1, Application US/10158865			
: Patent No. 6846651			
: GENERAL INFORMATION:			
: APPLICANT: Fleischmann et al.			
: TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fra			
: Patent No. 6846651			
: TITLE OF INVENTION: Thereof, and Uses Thereof			
: FILE REFERENCE: PB186P2C1D1			
: CURRENT APPLICATION NUMBER: US/10/158,865			
: CURRENT FILING DATE: 2002-06-03			
: PRIOR APPLICATION NUMBER: US 09/557,884			
: PRIOR FILING DATE: 2000-04-25			
: PRIOR APPLICATION NUMBER: US 08/476,102			
: PRIOR FILING DATE: 1995-06-07			
: PRIOR APPLICATION NUMBER: US 08/426,787			
: PRIOR FILING DATE: 1995-04-21			
: NUMBER OF SEQ ID NOS: 1			
: SOFTWARE: PatentIn version 3.1			
: SEQ ID NO 1			
: LENGTH: 1830121			
: TYPE: DNA			
: ORGANISM: Haemophilus influenzae			
: FEATURE:			
: NAME/KEY: misc_feature			
: LOCATION: (4747)..(4747)			
: OTHER INFORMATION: n equals a,t,c, or g			
: FEATURE:			
: NAME/KEY: misc_feature			
: LOCATION: (9921)..(9921)			
: OTHER INFORMATION: n equals a,t,c, or g			
: FEATURE:			
: NAME/KEY: misc_feature			
: LOCATION: (10150)..(10150)			
: OTHER INFORMATION: n equals a,t,c, or g			
: FEATURE:			
: NAME/KEY: misc_feature			
: LOCATION: (29298)..(29298)			
: OTHER INFORMATION: n equals a,t,c, or g			
: FEATURE:			
: NAME/KEY: misc_feature			
: LOCATION: (36543)..(36543)			
: OTHER INFORMATION: n equals a,t,c, or g			
: FEATURE:			
: NAME/KEY: misc_feature			
: LOCATION: (36551)..(36551)			
: OTHER INFORMATION: n equals a,t,c, or g			
: FEATURE:			
: NAME/KEY: misc_feature			
: LOCATION: (36636)..(36636)			
: OTHER INFORMATION: n equals a,t,c, or g			
: FEATURE:			
: NAME/KEY: misc_feature			
: LOCATION: (40808)..(40810)			
: OTHER INFORMATION: n equals a,t,c, or g			

1	FEATURE:	
2	NAME/KEY: misc_feature	
3	LOCATION: (44416)..(44416)	
4	OTHER_INFORMATION: n equals a,t,c, or g	
5	FEATURE:	
6	NAME/KEY: misc_feature	
7	LOCATION: (44905)..(44905)	
8	OTHER_INFORMATION: n equals a,t,c, or g	
9	FEATURE:	
10	NAME/KEY: misc_feature	
11	LOCATION: (44975)..(44975)	
12	OTHER_INFORMATION: n equals a,t,c, or g	
13	FEATURE:	
14	NAME/KEY: misc_feature	
15	LOCATION: (45593)..(45593)	
16	OTHER_INFORMATION: n equals a,t,c, or g	
17	FEATURE:	
18	NAME/KEY: misc_feature	
19	LOCATION: (45732)..(45732)	
20	OTHER_INFORMATION: n equals a,t,c, or g	
21	FEATURE:	
22	NAME/KEY: misc_feature	
23	LOCATION: (47036)..(47036)	
24	OTHER_INFORMATION: n equals a,t,c, or g	
25	FEATURE:	
26	NAME/KEY: misc_feature	
27	LOCATION: (51334)..(51334)	
28	OTHER_INFORMATION: n equals a,t,c, or g	
29	FEATURE:	
30	NAME/KEY: misc_feature	
31	LOCATION: (51602)..(51602)	
32	OTHER_INFORMATION: n equals a,t,c, or g	
33	FEATURE:	
34	NAME/KEY: misc_feature	
35	LOCATION: (51786)..(51786)	
36	OTHER_INFORMATION: n equals a,t,c, or g	
37	FEATURE:	
38	NAME/KEY: misc_feature	
39	LOCATION: (51805)..(51805)	
40	OTHER_INFORMATION: n equals a,t,c, or g	
41	FEATURE:	
42	NAME/KEY: misc_feature	
43	LOCATION: (53369)..(53369)	
44	OTHER_INFORMATION: n equals a,t,c, or g	
45	FEATURE:	
46	NAME/KEY: misc_feature	
47	LOCATION: (65309)..(65309)	
48	OTHER_INFORMATION: n equals a,t,c, or g	
49	FEATURE:	
50	NAME/KEY: misc_feature	
51	LOCATION: (65313)..(65313)	
52	OTHER_INFORMATION: n equals a,t,c, or g	
53	FEATURE:	
54	NAME/KEY: misc_feature	
55	LOCATION: (80024)..(80024)	
56	OTHER_INFORMATION: n equals a,t,c, or g	
57	FEATURE:	
58	NAME/KEY: misc_feature	
59	LOCATION: (100091)..(100091)	
60	OTHER_INFORMATION: n equals a,t,c, or g	
61	FEATURE:	
62	NAME/KEY: misc_feature	
63	LOCATION: (102696)..(102696)	
64	OTHER_INFORMATION: n equals a,t,c, or g	
65	FEATURE:	
66	NAME/KEY: misc_feature	
67	LOCATION: (105121)..(105121)	
68	OTHER_INFORMATION: n equals a,t,c, or g	
69	FEATURE:	
70	NAME/KEY: misc_feature	
71	LOCATION: (107248)..(107248)	
72	OTHER_INFORMATION: n equals a,t,c, or g	
73	FEATURE:	

```
NAME/KEY: misc_feature
LOCATION: (117136) .. (117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750) .. (119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924) .. (119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038) .. (120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344) .. (121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167) .. (122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336) .. (122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340) .. (131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360) .. (131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910) .. (139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398) .. (140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750) .. (142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058) .. (145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171) .. (145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942) .. (145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197) .. (147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841) .. (150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500) .. (152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
```

```
Query Match 66.7%; Score 971.8; DB 3; Length 1830121;
Best Local Similarity 81.3%; Pred. No. 0; Mismatches 267; Indels 7; Gaps 5;
Matches 1189; Conservative 0;

QY 1 ATTGAAGCTGGCGGATGCTTTATACATGCAAGTGAACGGGACGCG--ATGCTTGC 58
DB 128095 ATTGAAGCTGGCGGATGCTTTATACATGCAAGTGAACGGGACGCGGAAAGCTTGC 128036
QY 59 ATC--TGATGGCGAGTGGCGGACGGGTGATTAATGCAATCGGAACGTAATCCAGAAAGAGGG 116
DB 128035 TTTCTTGCTGACGAGTGGCGGACGGGTGATTAATGCTTGGGAATCTGCTTAATGAGGGG 127976
QY 117 GGTAAAGCATGCAAAATGCTTAATACCGCAATATCTCTAAAGAGGAAACGAGGGATC 176
DB 127975 GATTAACGACGGAAACCTGCTTAATACCGCTTAATATCGGAAGATGAAGAGTGGGACT 127916
QY 177 GAAAGACCTTGCGCTTTTGGAGCGCGGATGCTGAATTAAGCTAAGTTGGTGGGTTAAAGGC 236
DB 127915 GAGAGCGCGCATGCTCATAGATAGACCCAAAGTGGATTAAGTTGGTGGGTTAAATGC 127856
QY 237 CTACCAAGGCGACATCAATGATTTGGTCTGAGAGACGACCAAGCACAATGGGACTGAGA 296
DB 127855 CTACCAAGGCGATCTTAAGCTGCTGAGAGATGACCAAGCACAATGGGACTGAGA 127796
QY 297 CAGGCGCCAGATCTCTACGGGAGGACAGTGGGGAATTTTGGACAATGGGCGCAAGCT 356
DB 127795 CAGGCTCAGATCTCTACGGGAGGACAGTGGGGAATTTTGGACAATGGGCGCAAGCT 127736
QY 357 GATTCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGAAGCTCTTTCAGTGCAG 416
DB 127735 GAGCGACCATGCGCGGTGAGTGAAGAGGCTTGGGTTGAAGCTCTTTCAGTGCAG 127676
QY 417 AAGAAAAGTTACGGTAAATTAATGTCGACCCATGACGGTATCGACAGAGAGACCGGC 476
DB 127675 AAGAAAGTTAGTATGTAATAGTACATCAATGACGCTTAATACAGAGAGACCGGC 127616
QY 477 TAACCTACGTCAGACGACCGCGGTAATGCTAGGGTCCAGCGCTTAATCGGAATTAATCG 536
DB 127615 TAACTCCGTCAGACGACCGCGGTAATGCTAGGGTCCAGCGCTTAATCGGAATTAATCG 127556
QY 537 GCGTAAAGGGTGGCGACGCGGCTTGTAAATGAGTGAATATCCCGGCTTAATCGG 596
DB 127555 GCGTAAAGGGTGGCGACGCGGCTTGTAAATGAGTGAATATCCCGGCTTAATCGG 127496
QY 597 GAATTCGCTTGAATCACTAAGCTAGAGTGTGACAGAGGAGGTGAATTCATGTGTA 656
DB 127495 GAATTCGCTTGAATCACTAAGCTAGAGTGTGACAGAGGAGGTGAATTCATGTGTA 127436
QY 657 GCAGTGAATGCTAGAGATATGGAAGAATCATGATGCCAAGGACAGCTCTCTGGTTAA 716
DB 127435 GCGGTGAATGCTAGAGATATGGAAGAATCATGATGCCAAGGACAGCTCTCTGGTTAA 127376
QY 717 CACTGACGCTCATCAAGAAAGCGTGGGAGCAACAGATTAATACCTGTAGTCA 776
DB 127375 TACTGACGCTCATGCGAAGGCGTGGGAGCAACAGATTAATACCTGTAGTCA 127316
QY 777 CGCCCTAAACGATGTCACTAATGTTGTGGGCTTAATTAAGCTTGTGAACGAAGCTAACGC 836
DB 127315 CGCTGTAAACGCTGTCAATTTGGGATTTGGGCTTA--GAGCTTGGGCCGTAGCTTAACGT 127257
QY 837 GTGAAGTTGACCGCTTGGGAGTACCGTCCGAAGTTTAAACTCAAGAGAAATTTGACGGGG 896
DB 127256 GATAAATGCAACGCGCTGGGAGTACCGGTCGAAGTTTAAACTCAAGAGAAATTTGACGGGG 127197
QY 897 ACCGCAACAGCGGTGAATTAATGGAATTAATGATGCAACGCAAGAAACCTTAACCTAC 956
DB 127196 GCCGCAACAGCGGTGAATTAATGGAATTAATGATGCAACGCAAGAAACCTTAACCTAC 127137
QY 957 CTTGACATGTGAGGAATTTTCTAGAGAT--AGATTAAGTCTTGGGAAACGCTAACACAGG 1015
DB 127136 TCTTGACATCTTAAGAAAGAGCTCAGAGATGAGCTTGTGCGGAACTTAAGAGACAGG 127077
```



Oy	976	TTCTTAGAGATATGATTAGG-CTTCGGGAAACCTTAACAGAGTGTGATGGCTGTCTGCA	1034
Db	954	TTGCAGAGATCCGAGAGTGCCTTCGGGAATTCACTACAGGTGTCTGCATGGCTGTCTGCA	1013
Oy	1035	GCTCGTGTCTGTAGAGATGTTGGGTTAAGTCCCGCAACAGCGCAACCTTGTCTAATTAATGG	1094
Db	1014	GCTCGTGTCTGTAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGTCTTAAGTTA	1073
Oy	1095	CCATTC-ATTGGTTGGGCACTTTTAATGAGACTGCCCGTGCAAAACCGGAGGAAGGTGGGG	1153
Db	1074	CCAGCGACTCGGTGGGAACTTAAAGATCTGCAATGAACTGGAGAAAGCGGCGGG	1133
Oy	1154	ATGACGTCAAGTCCCTCATGAGCCCTTAATGAGGCTTCACACGTAAATACATGCGCGCT	1213
Db	1134	ACGACGTCAAGTCAATCATATGCGCCCTTACGACAGGGCTACACACGTGTCAATAGGTTGGT	1193
Oy	1214	ACAGAGGGTTGCCAACCAGCGGAGGAGAGCTAATCTCAGAAACCGCGTGTAGTCCGAT	1273
Db	1194	ACAAAGGGTTGCTACACAGCGATGAGTGTAAATCTCAAAAAGCCAATCGTAGTCCGAT	1253
Oy	1274	CGGATCTTGCAACTGCACTCCGTGAAGTCGGAATGCGTAATGCGGGATCAGCATGTC	1333
Db	1254	TGGAATCTGCAACTGCACTCCAGTAAGTCGGATGTCTAAGTATGCAATCAGATGCT	1313
Oy	1334	GCGGTGAATTAAGCTTCCCGGGTCTTTGACACACCGCCGCTCACACCATGGAGTGGGTTTC	1393
Db	1314	GCGGTGAATTAAGCTTCCCGGGCTTTGTACACACCGCCGCTCACACCATGGAGTTGATCTC	1373
Oy	1394	ACCAGAGCAGGTAGTCTAACCGTAAAGAGGGCGCTTGCACAGGTGAATTCATGACTGG	1453
Db	1374	ACCAAGAGGTGTTAAGCTTAAC--GCAGAGGGCGGATCACACAGTGGGGTGGATGAATCGG	1433
Oy	1454	GGTG 1457	
Db	1432	GGTG 1435	
RESULT 58			
US-10-278-942-1			
; Sequence 1, Application US/10278942			
; Patent No. 6720419			
; GENERAL INFORMATION:			
; APPLICANT: TAKARA BIO INC.			
; TITLE OF INVENTION: Sulfated fucan oligosaccharide			
; FILE REFERENCE: SAKX112			
; CURRENT APPLICATION NUMBER: US/10/278, 942			
; CURRENT FILING DATE: 2002-10-23			
; PRIOR APPLICATION NUMBER: JP 2001-325960			
; PRIOR FILING DATE: 2001-10-24			
; NUMBER OF SEQ ID NOS: 1			
; SEQ ID NO 1			
; LENGTH: 1506			
; TYPE: DNA			
; ORGANISM: Fucanobacter lyticus			
US-10-278-942-1			

Query Match	66.2%;	Score 964.6;	DB 3;	Length 1506;			
Best Local Similarity	81.3%;	Pred. No. 0;					
Matches 1193;	Conservative	0;	Mismatches 264;	Indels 10;			
			Gaps	6			
QY	1	ATTGAA	CGTGGCGG	CATCTTTAC	CATGCAAGTCGAACGGCAGACGAGATGCTTG	CAT	60
Db	21	ATTGAA	CGCTGGCGG	CAGGCTTAA	CAACATGCAAGTCGACGGAAAGAAATTA	AGCTTGCT	80
QY	61	CT--	GGTGG	CGAGTGGCGG	ACGGGTAGTAATTAATGCATCGA	ACGTATCC	118
Db	81	ATTGGC	GTGAGACGGCGG	ACGGGTAGTAATTAATGCATTAATG	CTTAATGCTTGGAGGA		140
QY	119	TAA	CGCATGAAAGATG	TGCTTAATACCGGATTA	TACTCTTAAGGAGGAAACAGAGGAT	CGA	178
Db	141	CAAC	AGTTGAAACGAC	TGCTTAATTAACGATTAATG	CTTACGGACCAAAAGAGGGGAT	TCT	200

Qy	179	AAG--ACCTTCGCGTTTGGAGCGCGCGAAGTCTGAAATTGAGTATGTGGGGTTAAAGG	236
Db	201	TGGAAACCTTTCGGCAATTTGATTAAGCCCAAGTAGAAGTATGCTATGTAAGGTAAATGGC	260
Qy	237	CTACCAAGGCGAGCATCAGTAGTATGGTCTAGAGAGACGACCAAGCCCACTGGAGCTGAGA	296
Db	261	TTACTCTAGGCAACATCTTAAGCTGGTTTGAAGAGATGATCAACCCACACTGGGACTGAGA	320
Qy	297	CACGGCCCAACTCTTACGCGGAGGCAAGCACTGGGGAAATTTTGGACAAATGGGCCAAGCCT	356
Db	321	CACGGCCCAAGCTCTTAAGGGAGGCAAGCTGGGGAAATATTTGACAAATGGGCCAAGCCT	380
Qy	357	GATTCAGCAATCCCGCGTAGTGAAGAAAGCCTTCGGGTTGTAAAGCTCTTTCAATCGAG	416
Db	381	GATGACAGCAATCCCGCGTGTGAAGAAAGCCTTCGGGTTGTAAAGCACTTTACGCGAGG	440
Qy	417	AAGAAAAGGTAAAGGTAAATTAATGCTGACCACTAGCGGTATCGACAGAAAGAACCGGCG	476
Db	441	AAGAAAAGGTATTAATTAATCTTGCAATCTGTGACCTTACTGCGAAGAAAGAACCGGCG	500
Qy	477	TAACTACGTCGACAGCGCGGCTAATACGTAGGGTCAAGCGTTAAATCGAAATTAATCTGG	536
Db	501	TAACTCTGCGACAGACGCGGCTAATACGAGGGGTCAAGCGTTAAATCGAAATTAATCTGG	560
Qy	537	GCGTAAAGGGTCCGAGGCGCGCTTTGTAAGTCAATGTGAATATCCCGGGCTTAACCTGG	596
Db	561	GCGTAAACGTCGCTAAGGTGTGTTTAAGCAATATGTGAAGGCCCGGGCTCAACCTGG	620
Qy	597	GAAATTGCGTTGAAACCTACAAAGCTAGAGTGTGACAGAGGAGGTGAATTTCCATGTGTA	656
Db	621	GAACTGCATTTTGAATCTGGCAAACTAAGATTTTGAAGGGTAGTGAATTTCCAGTGTA	680
Qy	657	GCAGTGAATCCGTAGAGATATGGAAGAACATCGATGGCGAAGCGCAAGCCTTCGGGTTAA	716
Db	681	GCGGTGAATCCGTAGAGATATGGAAGAACATCACTGGCGAAGCGGCTACTGAGACAGA	740
Qy	717	CAGTACGCTCATGCAAGAAACCGTGGGAGCAAAACAGATTAAATACCTCTGTAGTCCA	776
Db	741	GACTGACACTGAGGCAAGAAACCGTGGGAGCAAAACAGATTAAATACCTCTGTAGTCCA	800
Qy	777	CGCCCTTAACAGATGTCATCTAGTTGT--TGCGCTTAATTAAGCTTGGTATGACAACTAACG	835
Db	801	CGCCCTTAACAGATGTCATCTAGCGCTCTGTAGACTTGAATCTGTGGGTGGCTTAAGC	860
Qy	836	CGTGAAGTTGACCGCGCTGGGAGTACCGTTCGCAAGATTAAACTCAAGAAATTGACGGG	895
Db	861	CGCTAAGTTGACCGCGCTGGGAGTACCGCGCCAAAGTTAAACTCAATGAATTGACGGG	920
Qy	896	GACCCGCAACAGCGGTGATTAATGTGATTAATTCGATGCAACCGCAAAAACCTTAACCTA	955
Db	921	GGCCCCGCAACAGCGGTGAGATGTTTAATTCGATGCAACCGCAAAAACCTTAACCAT	980
Qy	956	CCCTTGAATGTAGGAATTTTCTAGAGATTAAGT--CTTGGGGAAGGCTAAACAG	1014
Db	981	CCCTTGAATCTCTACTTAAGTTACTAGAAATGTTTCGCTTGGGGAAGTAAGTACAG	1040
Qy	1015	GTGCTGCAATGGCTGTCAGCTCGTGTGTAAGATGTTGGGTTAAAGTCCCGCAACGAGC	1074
Db	1041	GTGCTGCAATGGCTGTCAGCTCGTGTGTAAGATGTTGGGTTAAAGTCCCGCAACGAGC	1100
Qy	1075	GCAACCCCTTGATTAAATATGTCATATTTGGTTGGCG--ACTTTAATGAGACTGCGCGTG	1132
Db	1101	GCAACCCCTTACTTAATTTGCTAGCGGCTAAATGGCGAAGACTTAAGAGACTGCGCGTG	1160
Qy	1133	ACAAACCCGAAGGAAGTGGGAGTACGATCCTCATGGCCCTTAATGGTAAAGGCTTC	1192
Db	1161	ATPAACCCGAAGGAAGTGGGAGCGAGCTCAAGTCTCATATGGCCCTTAAGGGAATGGCTTAC	1220
Qy	1193	ACACGTAAATCAATGGCGCTGACAGAGGGTTTCCAAACCCGAGGAGGGAGCTAATCTTACG	1252
Db	1221	ACACGTGCTACAAATGGCAAGTACAGAGGGCAACAATACCGGAGGTGGAGGAATTCACAC	1280
Qy	1253	AAAGCGCTGTAGTCCGAGTCCGATCTGCAATCTCACTCCGTGAAGTCCGAATTCGCTA	1312

Db	1281	AAAGCTTGTCCTAGCCGATTTGGAGTGTGCACACTCGACTCATGAATCGAATCGCTA	1340
Qy	1313	GTAATCGGGATCAGCAATGTGGCGGTGAAATACGTTCCCGGGGTCTTTGACAACGCCCGT	1372
Db	1341	GTAATCGTAGATCGAAGAATGCTAACGGTGAATACGTTCCCGGGCTTTGTACACACGCCCGT	1400
Qy	1373	CACACCATTGGAGTAGGGTTTACCAACGAGACGGTAGTCTAACCGT--AAGAGGGGCGCTT	1430
Db	1401	CACACCATTGGAGTAGGGTTGCAAAAGAGTGGCTAATTTAACCTTCGGGAGGACGGTC	1460
Qy	1431	GCCACGGTAGATTCATGACTGGGGTG	1457
Db	1461	ACCACCTTTGTATTCATGACTGGGGTG	1487

```

RESULT 59
US-10-694-352-1
; Sequence 1, Application US/10694352
; Patent No. 6927289
; GENERAL INFORMATION:
; APPLICANT: TAKAKA BIO INC.
; TITLE OF INVENTION: Sulfated fucan oligosaccharide
; FILE REFERENCE: SAV4112
; CURRENT APPLICATION NUMBER: US/10/694,352
; CURRENT FILING DATE: 2003-10-28
; PRIOR APPLICATION NUMBER: US/10/278,942
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: JP 2001-325960
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Fucanobacter lyticus
US-10-694-352-1

```

Query Match	Similarity	Score	964.6	DB 3	Length	1506			
Best Local	Similarity	81.3%	Pred.	No. 0					
Matches	1193	Conservative	0	Med. Matches	264	Indels	10	Gaps	6
QY	1	ATTGACGCTGGCGGCGATGCTTTTACATATGCATGCAAGTCGAAACGGGACGACGGATGCTTGATC	60						
Db	21	ATTGAACGCTGGCGGCGAGCTTTAACATATGCATGCAAGTCGAGCGGAAACGAGATATAGCTTGCT	80						
QY	61	CT--CGTGGCAGATGGCCGACGGGTGAGTAATGCAATCGGAACGTATCCAGAAAGAGGGGGG	118						
Db	81	ATTGGGCGTCGAGCGGCGGACGGGTGAGTAATGCTTGGGAATATAGCTTATATGTGTGGGGA	140						
QY	119	TAAACGATCGAAAGATGTGCTAATACCGCATATACTCTAAGGAGGAAAGCAGGGATCGA	178						
Db	141	CACAGTTGGAAACGACTGCTAATACCGCATATATGCTATCGGACCAAGAAAGAGGGGATTC	200						
QY	179	AAG--ACCTTGCGCTTTTGGAGCGGCGCATGTCTATTAATGCTAATTTGGTGGGTAAAGC	236						
Db	201	TGGAACCTTTGCGCAATTTGATTAACCCCAAGGAGATTAAGCTAATGATTAAGTATATGTC	260						
QY	237	CTACCAAGGCGACGATCTAGTAGTTGGTCTGAAAGACGACACGACACTGGGACTGAGA	296						
Db	261	TTACTTAGGCGACGATCTCTAGCTGGTTGGAAGAGATGATCAGCCACTGGGACTGAGA	320						
QY	297	CACGGCCGAGACTCTTAAGGAGGCGACAGTGGGGAATTTTGGACATGGGCGCAAGCTT	356						
Db	321	CACGGCCGAGACTCTTAAGGAGGCGACAGTGGGGAATTTTGCACAAATGGGGGAAAGCTT	380						
QY	357	GATTCAGCAATGCGCGGTGAGTGAAGAAAGGCGTTGGGGTGTAAAGCTTTTCAGTCGAG	416						
Db	381	GATGAGCGCATGCCCGCTGTGTGAAGAAAGGCTTTGGGGTGTAAAGCACTTTCAAGCAGG	440						
QY	417	AAGAAAGGTTACGGTAAATATCTGTGACCATGACGGTATCGACAAAGAAAGCACCGGC	476						
Db	441	AGGAAGGGTGTAGTTAATATCTGTGCATCTGTGACGTTACTTCGGAGAGAAAGCACCGGC	500						

QY	477	TAAC	TACGTCGACGACACCCGCGTAATACGTAAGGGTGCACACGCTTAATCCGAATTACTCG	536
Db	501	TAAC	TCCTGTGTGCACACACCCGCGTAATACGTAAGGGTGCACACGCTTAATCCGAATTACTCG	560
QY	537	GCGTAAAGGGGCGCCAGCGCGCTTGTAGTCAGATGTGAATATCCCGGGCTTAACCTCG	596	
Db	561	GCGTAAAGCGTGGCTAGGTGGTTTGTTAAGCAGATGTGAAGGCCCGGGCTCAACCTCG	620	
QY	597	GAATTGGCTTTGAAACTACAAAGCTTAAGTGTGGCAGAGGGAGGTTGGAATTCATGTGTGA	656	
Db	621	GAAC	TGCAATTTTGAACCTCGCAAACTAAGTTTGTGAAGGGTATGGAATTTCCAGTGA	680
QY	657	GCAGTGAAGTCCGTAAGATATATGGAAGAACTCGATGTGCGCAAGGCACCTCTCGGTAA	716	
Db	681	GGGGTGAAGATCGTAAGATTTGAAGAGAACATCACTGTGCGAAGGCAGCTACTGTGACAGA	740	
QY	717	CAC	TGAGCTCATCGACGAAAGCCGTGGGGAGCAAAACAGATTAGATACCTGTAGTCCA	776
Db	741	GAC	TGACACTGAGGCAAGAAAGCGTGGGGAGCAAAACAGATTAGATACCTGTAGTCCA	800
QY	777	CGCCCTAAAGCATGTCAACTAATGTTGT-TGGG	835	
Db	801	CGCCGTAAACGATGTCAACTAGCCGCTGTAGACTTGATCTGTGGGGTGGCGTACTAAG	860	
QY	836	CGTGAAGTGAACCGCTGGGGAGTACGAGTGCAGAGTTTAAACTCAAGAGATTTGAACGG	895	
Db	861	CGCTAAGTTGACCGCTCGGGAGTACAGCCGCGCAAGTTTAACTCAATGAATTTGACGG	920	
QY	896	GACCCGCAACAGCGGTGGATTAATGTGATTAATTCATGCACCGCGAAACCTTACTTA	955	
Db	921	GGCCCGCAACAGCGGTGGAGATGTGGTTTAATTCATGCACCGCGAAACCTTACTTA	980	
QY	956	CCCTTGACATGTAGCGAATTTTCTAGAGATTAAGATTATGTG-CTTGGGAACGCTAACACAG	1014	
Db	981	CCCTTGACATCTTAAGTTACTAAGTAACTAAGATTAAGTTTGTGCTTCGGGAAGTACTGACAG	1040	
QY	1015	GTGCTGATGAGCTGTGTCAGAGCTGTGTCGTAAGATTTGGGTTAAATCCCGACAGAGC	1074	
Db	1041	GTGCTGATGAGCTGTGTCAGAGCTGTGTCGTAAGATTTGGGTTAAATCCCGACAGAGC	1100	
QY	1075	GCAACCCCTGTGATTAATGGCCATCTTTGGTGGGC--ACTTAAATGACATCGCCGCTG	1132	
Db	1101	GCAACCCCTTAATCTTATTTGCTACCGGTAATGCGAGAACTCTAAGAGACTCCCGGCTG	1160	
QY	1133	ACAAACCGGAGGAGTGTGGGAGTACGTCAAGTCTCATGGCCCTTAAATGGGTAGGGCTTC	1192	
Db	1161	ATTAACCGGAGGAGTGTGGGAGTACGTCAAGTCTCATGGCCCTTAAACGGGATGGGCTAC	1220	
QY	1193	ACAGCTAATACAAATGCGCGCTACAGAGGGTGTCCAAACCCCGAGGGGAGCTAATCTCAG	1252	
Db	1221	ACAGCTGCTAACAAATGCGCAAGTACAGAGGGGACGAATACCGGAGAGTGAAGCAATCCAC	1280	
QY	1253	AAAGCGCGTGTAGTGTCCGAGATCCGAGTCTGCACTCGATCTCTGTAAGTCCGAATTCGCTA	1312	
Db	1281	AAACCTTGCTGTAGTGTCCGAGATTTGAGTCTGCACTCGATCTCTCAATGAAGTCCGAATTCGCTA	1340	
QY	1313	GTAATCCGGAATCAAGCATGTGCGGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGCT	1372	
Db	1341	GTAATCCGAGATCAAGAAATGCTACGCTGTGAATACGTTCCCGGGCTTGTACACACCGCCGCT	1400	
QY	1373	CACACCAATGGAGAGTGGGTTTCCACCAAGACAGATAGTCTAACCGT--AAGAGGGCGCTT	1430	
Db	1401	CACACCAATGGAGAGTGGGTTTCCCAAAAGAGTGGCTAGTTTAACTCTTCGGGGAGAGCGGCTC	1460	
QY	1431	GCCACGGTAGATTCATGACTGGGGGTG	1457	
Db	1461	ACCACTTGTGATTCATGACTGGGGGTG	1487	

RESULT 60  
US-09-726-774-13  
; Sequence 13, Application US/09726774  
; Patent No. 6677153



```

; GENERAL INFORMATION:
; APPLICANT: Iversen, Patrick L.
; TITLE OF INVENTION: Antisense Antibacterial Method and
; FILE REFERENCE: 0450-0032.30
; CURRENT APPLICATION NUMBER: US/09/726,774
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 60/168,150
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Hemophilis influenza
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1487)
; OTHER INFORMATION: n = A,T,C or G
;
US-09-726-774-13

Query Match      66.1%; Score 963.8; DB 3; Length 1487;
Best Local Similarity 80.8%; Pred. No. 0;
Matches 1181; Conservative 0; Mismatches 273; Indels 7; Gaps 5;

QY      1  ATTGAACGCTGGCGGCAATGCTTTACACATGCAAGTCGACGCGCA--ATGCTTGC 58
DB      28  ATTGAACGCTGGCGGCGGCTTAACATGCAATGCAAGTCGACGCGCAAGAGCTTGC 87
QY      59  ATC--TGTGGCGAGTGGCGGAGCGGTGATATGCAATCGTAACGTATCCAGAAAGAGGG 116
DB      88  TTCTTCTCTGACGAGTGGCGGAGCGGTGATATGCTTGGGAAATCTGCTTATGAGAGGG 147
QY      117 GGTAAACCATGGAAGATGCTTAATACCGCATATCTCTAAGAGGAAAGACAGAGGATC 176
DB      148 GATTAACACAGGGAATCTGCTCTAATACCGCTATTATCGGAAGATGAAGTGGCGGACT 207
QY      177 GAAAGACCTTGGCGCTTTTGGAGCGCGCATGCTGATTTAGCTAGTGGTGGGTAAAGGC 236
DB      208 GAGAGCGCGCATGCGCATAGATGAGCCCAAGTGGGATTAAGTAGTGGGTAAATGC 267
QY      237 CTACCAAGCGCATCATGATGTTGTTGTTGAGAGACACACGCACTGGGACTGAGA 296
DB      268 CTACCAAGCGCTGATCTCTAGCTGCTGAGAGGATGACAGCCACTGGAACCTGAGA 327
QY      297 CACGCGCCAGACTCTCTACGCGGAGGAGAGTGGGGAAATTTTGAACAATGGGCGAAGCT 356
DB      328 CACGCTCCAGACTCTCTACGCGGAGGAGAGTGGGGAAATTTGCGNATGGGCGAAGCT 387
QY      357 GATCAGCAATGCGCGGAGTGAAGAGGCTTTCGGGTGTAAAGCTCTTTCAGTCCAG 416
DB      388 GAGCGAGCCATCGCGCGTGAATGAAGAGGCTTTCGGGTGTAAAGCTCTTTCAGTATG 447
QY      417 AAGAAAAGTTACGGTAAATATCGTGAACCATGACGGTATCGACAGAAAGACCGGC 476
DB      448 AGAAGAGTTAGTGTATATAGCACATCAATTAAGCTTTAATACAGAAAGACCGGC 507
QY      477 TAACTAGGTGCGACAGCGCGGCTTAATCTGAAGGTGCAACCTTAATCGGAATTAACG 536
DB      508 TAACCTCCGTGCGAGAGCGCGGTAAATCGAGAGTGGAGCGTTAATCGGAATTAACG 567
QY      537 GCGTAAAGGTTGCGGAGCGCGCTTGAATGATGTAATGTAATCCCGGGCTTAAACCTGG 596
DB      568 GCGTAAAGGTTGCGGAGCGCGCTTGAATGATGTAATGTAATCCCGGGCTTAAACCTGG 627
QY      597 GAATTCGCTTTGAACCTAACAAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGA 656
DB      628 GAAATTCGCTTTGAAGCTGGGTAACTTAAGAGGAGGTGAATTCATCAAGTGA 687
QY      657 GCAATGAATGCTGATGATATGAGAAACATGATGCGGAAGGAGGAGCTCTGGGTAA 716
DB      688 GCGGTGAATGCTGATGATGATGAGAAATCCGAAGGAGGAGGAGGAGCTCTGGGTAA 747
```

```

QY      717 CACTGACGCTATGACAGAAAGCGTGGGAGGAGAAACAGATTAAGATACCTGGTACTCA 776
DB      748 TACTGACGCTATGATGCGAAAGCGTGGGAGGAGAAACAGATTAAGATACCTGGTACTCA 807
QY      777 CGCCCTAAACGATGTCACATGTTGTTGGGCTTTATTAAGCTTGTGAAGAGCTAACGC 836
DB      808 CGCTGAAACGCTGCTGATTTGGGGGTGGGGGT--TAACTTGGCACCCTGACTAACGT 866
QY      837 GTAAAGTTACCGCCCTGGGAGTACCGGTGCAAGATTAATAACCAAGAAATTGACGGGG 896
DB      867 GATTAATGACCGCTGGGAGTACCGCGCAAGGTTAAACTCAATGATTAATGACGGGG 926
QY      897 ACCGCAAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 956
DB      927 GCGGCAAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 986
QY      957 CCTTGAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1015
DB      987 TCTTGAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1046
QY      1016 TGTGCAATGCTGTGCTGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
DB      1047 TGTGCAATGCTGTGCTGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
QY      1076 CAACCTTGTATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1134
DB      1107 CAACCTTGTATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1166
QY      1135 AAACCGGAGGAGTGGGAGTGAAGTCAAGTCTCTCAATGCTTATGAGGTAGGGCTTAC 1194
DB      1167 AAACCGGAGGAGTGGGAGTGAAGTCAAGTCTCTCAATGCTTATGAGGTAGGGCTTAC 1226
QY      1195 ACCTAATCAATGAGCGCTGACAGAGGTTCCAAACCGGAGGAGGAGTAACTCAGAA 1254
DB      1227 ACCTAATCAATGAGCGCTGACAGAGGTTCCAAACCGGAGGAGGAGTAACTCAGAA 1286
QY      1255 AGCGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
DB      1287 AGTACGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1346
QY      1315 AATGCGGATCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1374
DB      1347 AATGCGGATCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1406
QY      1375 CACCATGAGAGTGGGTTTACCAAGAGCAGTATGCTTAACCTGAAGAGGCGCTTGCCA 1434
DB      1407 CACCATGAGAGTGGGTTTACCAAGAGTATAGCTTAACCTTGGAGGGCGTTTACCA 1466
QY      1435 CGGTGAGATTCATGACTGGGG 1455
DB      1467 CGGTGAGATTCATGACTGGGG 1487

RESULT 61
US-09-726-774-1
; Sequence 1, Application US/09726774
; Patent No. 6677153
; GENERAL INFORMATION:
; APPLICANT: Iversen, Patrick L.
; TITLE OF INVENTION: Antisense Antibacterial Method and
; FILE REFERENCE: 0450-0032.30
; CURRENT APPLICATION NUMBER: US/09/726,774
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 60/168,150
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
```

```

; NAME/KEY: m1sc_feature
; LOCATION: (1)...(1450)
; OTHER INFORMATION: n = A,T,C or G
US-09-726-774-1

Query Match      64.9%; Score 945.6; DB 3; Length 1450;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1156; Conservative 0; Mismatches 257; Indels 8; Gaps 5;

QY 1 ATTGAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGACGAGATCCTTGCA- 59
   |||||
DB 19 ATTGAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 78
   |||||

QY 60 ---TCTGTGCGAGATGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 116
   |||||
DB 79 TCTGTGCGAGATGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 138
   |||||

QY 117 GGTAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 176
   |||||
DB 139 GATTAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 198
   |||||

QY 177 GAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 236
   |||||
DB 199 TTAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 258
   |||||

QY 237 TTAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 296
   |||||
DB 259 TCACTGCGAGATGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 318
   |||||

QY 297 CAGGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 356
   |||||
DB 319 CAGGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 378
   |||||

QY 357 GATCAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 416
   |||||
DB 379 GATCAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 438
   |||||

QY 417 AAGAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 476
   |||||
DB 439 AAGAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 498
   |||||

QY 477 TTAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 536
   |||||
DB 499 TTAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 558
   |||||

QY 537 GCGTAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 596
   |||||
DB 559 GCGTAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 618
   |||||

QY 597 GAAATGCGTAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 656
   |||||
DB 619 GAAATGCGTAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 678
   |||||

QY 657 GCGTAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 716
   |||||
DB 679 GCGTAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 738
   |||||

QY 717 CACTGAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 776
   |||||
DB 739 GACTGAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 798
   |||||

QY 777 GCGCTTAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 835
   |||||
DB 799 GCGCTTAAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 858
   |||||

QY 836 CGTGAAGTGAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 895
   |||||
DB 859 CGTGAAGTGAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 918
   |||||

QY 896 GACCGGCAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 955
   |||||
DB 919 G-GCCGCAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 977
   |||||

```

```

QY 956 CCCTGACATGTAAGCAATTTCTAGAGT-AGATTAGTCTTCGGGAACGTTAACACAG 1014
   |||||
DB 978 GTCTGACATGTAAGCAATTTCTAGAGT-AGATTAGTCTTCGGGAACGTTAACACAG 1037
   |||||

QY 1015 GTCTGACATGTAAGCAATTTCTAGAGT-AGATTAGTCTTCGGGAACGTTAACACAG 1074
   |||||
DB 1038 GTCTGACATGTAAGCAATTTCTAGAGT-AGATTAGTCTTCGGGAACGTTAACACAG 1097
   |||||

QY 1075 GCAACCTTTCATTAATTCGATC-ATTGTTGGGCACTTTAATGAGTCCGAGTCA 1133
   |||||
DB 1098 GCAACCTTTCATTAATTCGATC-ATTGTTGGGCACTTTAATGAGTCCGAGTCA 1157
   |||||

QY 1134 CAACCGGAGAGTGGGAGATGAGTCAAGTCTTCATGCTTCATGAGTTCGAGTCA 1193
   |||||
DB 1158 TAACTGAGAGAGTGGGAGATGAGTCAAGTCTTCATGCTTCATGAGTTCGAGTCA 1217
   |||||

QY 1194 CACGTAAATCAATGCGCGCATGCAAGAGTTCGCAACCGCGGAGGAGTCAATCTCAGA 1253
   |||||
DB 1218 CACGTAAATCAATGCGCGCATGCAAGAGTTCGCAACCGCGGAGGAGTCAATCTCAGA 1277
   |||||

QY 1254 AAGCGGTGCTAGTCCGAGTTCGAGTTCGCAACCGCGGAGGAGTCAATCTCAGA 1313
   |||||
DB 1278 AAGCGGTGCTAGTCCGAGTTCGAGTTCGCAACCGCGGAGGAGTCAATCTCAGA 1337
   |||||

QY 1314 TAACTGAGAGAGTGGGAGATGAGTCAAGTCTTCATGCTTCATGAGTTCGAGTCA 1373
   |||||
DB 1338 TAACTGAGAGAGTGGGAGATGAGTCAAGTCTTCATGCTTCATGAGTTCGAGTCA 1397
   |||||

QY 1374 ACACATGAGAGTGGGAGTTCACCAAGAGTCAATCTCAGA 1414
   |||||
DB 1398 ACACATGAGAGTGGGAGTTCACCAAGAGTCAATCTCAGA 1438
   |||||

RESULT 62
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGEMOBU, SHUJI
; APPLICANT: WATANABE, HIDEKI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match      64.6%; Score 941.2; DB 3; Length 640681;
Best Local Similarity 79.9%; Pred. No. 1,86313;
Matches 1174; Conservative 0; Mismatches 283; Indels 13; Gaps 5;

QY 1 ATTGAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 54
   |||||
DB 274064 ATTGAAGCTGCGCGCATGCTTTTACATGCAAGTCGAAACGAGATCCTTGCA- 274123
   |||||

QY 55 TTGCATCTGCTGCGAGTGGCGGACGCGGTGAGTATGCAATCGAACTGATCCGAAGAG 114
   |||||
DB 274124 CTCTCTTTGCGGCAAGCGGCAAAACGCGTGAATATCTGCGGATCTACCCAAAGAG 274183
   |||||

QY 115 GCGGTAAAGCTGCAAGATGCTTAATACCGATATATCTTAAGAGAGAGAGAGAGAG 174
   |||||
DB 274184 GCGGTAAAGCTGCAAGATGCTTAATACCGATATATCTTAAGAGAGAGAGAGAGAG 274243
   |||||

QY 175 TCGAAAGACCTTTCGCTTTTGAAGCGCGCATGCTTAAGTCTGATGCTGCTGCTTAAG 234
   |||||

```

Db	274244	CTTTTGGCGCTCATCGCTCTTTTGGATGAAACCCAGACGAGATTAGCTTGGTAGAATAA	2743033
Qy	225	GCCTAACAAAGGCGAGATCAGTAGTGGTCTGAGAGAGACAGCAGCACACTGGAGCTGA	294
Db	274304	GCTTACCAAGGAGAGATCTTAGCTGGTCTGAGAGAGATTAACAGCCACACTGGAACTGA	274353
Qy	295	GACACGGCCACAGACTCTTACGGGAGGACAGCAGTGGGAAATTTTGGCAATGGCGCGAAC	354
Db	274384	GACACGGTCCAGACTCTTACGGGAGGACAGCAGTGGGAAATTTTGGCAATGGCGCGAAC	2744233
Qy	355	CTGATCCAGCAATGCCCGCTGAGTGAAGAGCCTTCCGGTGTAAAGCTCTTCACTG	414
Db	274424	CTGATCCAGCTATGCCCGCTGATGAAAGAGCCTTACGGTGTAAAGTCTTCAAGCG	274483
Qy	415	AGAGAAAAGGTTACGGTAAATATCTGACCCATGACGGTATCGACAGAGAAAGACCG	474
Db	274484	AGAGAAAAGGTTAAATCTAATTAATTTTCTGTACCTTACCCGAGAGAGACCG	2745433
Qy	475	GCTACTACGTGCCAGACAGCCCGGTAAATACGTAGGGTCCAAAGCCTTAATCGAATTA	534
Db	274544	GCTTACTCGGTCCAGACAGCCCGGTAAATACGTAGGGTCCAAAGCCTTAATCAGAA	2746033
Qy	535	GGGCGTAAAGGCTGCGCAGCGCGCTTGTATGATCAGATGTGAAATCCCGGGCTTA	594
Db	274604	GGGCGTAAAGACCGCGCTAGTGTCTTTTAAAGTCAAGTGTGAAATCCCTAGGCTCA	2746633
Qy	595	GGGAATTTGGCTTGAACATCAAAAGCTAGAGTGGCAGAGGAGGAGTTCATGTG	654
Db	274664	AGGAATCTGATTTGAACATCGAATACTAGATGGCGAAGCGCTTCTGGCTT	2747233
Qy	655	TAGCAGTGAATAGCGTAGAGATATGGAAGAACATCGATGGCGAAGCGCTTCTGGCTT	714
Db	274724	TAGCGGTGAATAGCGTAGATATCGAAGGAATCTGAGAGGAATACCGGTGGCAAGCGCTT	2747833
Qy	715	AACTGACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAATACCTCTGTATG	774
Db	274784	AAACTGACATGAGCGCGGAAAGCGTGGGAGCAAAACAGATTAATACCTCTGTATG	2748433
Qy	775	CACGCGCTTAAAGATGTCAACTAGTTGGTGGCCTTATTAGGCTTG- GTACGAAAGTAA	833
Db	274844	CATGCGGTAAAGATGTCAACTAGTTGGTGGTCTTCCAGAGAAAGTACTTCCAAAGTAA	2749033
Qy	834	CGCGTGAAGTGAACCGCTGGGGAGTAGACGTCGCAAGATTTAAACTCAAGAAATTGAC	893
Db	274904	CGCATTAATTCGACCGCTGGGGAGTAGACCGCGCAAGGCTTAAACTCAATGAATTGAC	2749633
Qy	894	GGGACCCCGCAACAGCGGTGATTTATGTGATTAATTCGATCAACCGGAAAACTTTAC	953
Db	274964	GGGCGCCCGCAACAGCGGTGATGTGTTTAACTGATGCAACCGGAAAACTTTAC	2750233
Qy	954	TACCTTGAATGTAGCGAATTTTCTAGAGATAGATTAGT-CTTGGGAGCGTAAAC	1012
Db	275024	TGGTCTTGAATGTAGCGAATTTTCTAGAGATAGATTAGT-CTTGGGAGCTGTGAC	2750833
Qy	1013	AGGTGCTGATGCTGTGCTCAGCTCGTGTCTGAGAGATGTTGGGTTAAGTCCGCAACGA	1072
Db	275084	AGGTGCTGATGCTGTGCTCAGCTCGTGTGTGGAATGTTGGGTTAAGTCCGCAACGA	2751433
Qy	1073	GGGCAACCCCTGTCAATTAATTCGCAAT-ATTGGTGGGCACTTTAATGAGCTGCCGT	1131
Db	275144	GGGCAACCCCTTATCCCTGTGTGCAAGCGGTTCGGCGGGAACTCAGAGGAGATGCCGT	2752033
Qy	1132	GACAAACCGGAGAGAGTGGGAGTAGACGTAAGTCCATGAGCCCTTATGGTATGGGCTT	1191
Db	275204	TATTAACCGGAGAGAGTGGGAGTAGACGTAAGTCAATGAGCCCTTACGACAGGGCTA	2752633
Qy	1192	CACACGTAAATACATGGCGGTACAGAGGGTTCGCAACCGCGAGGGGAGTATCTCA	1251
Db	275264	CACACGTCTACATGGTCTTATTAAGAGAGAAATCTGCAAAAGCAAGCAAACTCA	2753233
Qy	1252	GAAAGCGCTGTATGTCGAGATCGAATCTGCACTCGATCCGTGAATGGAATGCT	1311

```

Db      275324  TAAATTAATCTAGTCCGAGCTGAGATGTGCACATCCGACTCCACGAAAGTCGGAATTCGCT 275383
Oy      1312    AGTATCCCGGATTCAGCATGTGCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCCG 1371
Db      275384  AGTATCGTGATTCGAATATGCACCGGTGAATACGTTCCCGGCGCTTTGTACACACCGCCCG 275443
Oy      1372    TCACACCATGGAGTGGGTTTCACCGAAAGCAGGTAGTCTAACCGT----AAGAGGGCG 1427
Db      275444  TCACACCATGGAGTGGGTTTCAACGAAAGCAGGTATCTTAAACCTTTTAAAGAGAGCGC 275503
Oy      1428    CTTGCGACGCTGAGATTCACTGACTGGGGTG 1457
Db      275504  CTTACCATTTGTGATTCATGACTGGGGTG 275533

RESULT 63
US-10-761-509-1
; Sequence 1, Application US/10761509
; Patent No. 6919197
; GENERAL INFORMATION:
; APPLICANT: Gerber, John F
; APPLICANT: White, James H
; TITLE OF INVENTION: Materials and Methods For The Efficient Production Of Pasteurized
; FILE REFERENCE: GER-100XC1
; CURRENT APPLICATION NUMBER: US/10/761,509
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US/09/636,185
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 60/148,154
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Helper Factor DNA is not Pasteuria but has 98% homology with
; OTHER INFORMATION: Enterobacter cloacae and Pantoea ssp. and others.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (99)_(99)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102)..(102)
; OTHER INFORMATION: n = a, c, g, or t.
US-10-761-509-1

Query Match      63.5%; Score 924.8; DB 3; Length 1455;
Best Local Similarity 82.8%; Pred. No. 1.9e-309;
Matches 1091; Conservative 0; Mismatches 224; Indels 3; Gaps 3;

Oy      60      TCTGTGTCGCGAGTGGCGGACCGGTGACTATGCATCGGAACGTATCCAGAGAGGGGGGT 119
Db      50      TTGTCGGGAGGCGGGGACGGGTGATATGTCGTGGAAATCTAGTGGANGATGGGGAT 109
Oy      120      AACGCATCGAAAGATGTGCTAATACGSCATATATCTTAAGGAGGAAAGCAGGGATCGAA 179
Db      110      CACTACTGAAACAGTGTCTAATACGSCATTAAGGTCTCAAGACCAAGAGGGGGGCTTTC 169
Oy      180      AGACCTTGGCGCTTTTGGAGCGCGCATGTCTGATTAGCTAGTGGTGGGTTAAAGGCTTA 239
Db      170      GGCGCTCTTGCCATCAGATGTGCCCAGATGGGATTACTGTAGTGGGGTTAACGGCTCA 229
Oy      240      CCAAGCGCAGCATCAGTAAGTTGTCGAGAGAGCAGCAGCAGCAGTGGGACTGAGACAC 299
Db      230      CTAAGCGCAGCATCCTTACGTGTGTCGAGAGATGACACAGCAGCAGTGGGAACTGAGAC 289
Oy      300      GGCCCAAGCTCTTACGCGAGGAGCAGCAGTGGGGAATTTTGAACAATGGCGGCAAGCTGAT 359
Db      290      GGTCGAGACTCTTACGCGAGGAGCAGCAGTGGGGAATTTTGCACATGGGCGCAAGCTGAT 349

```

QY 360 CCAGCAATGCCGCTGATGTAAGAAGGCTTTCGGGTTTGAAGCTTTCACTCGAGAG 419  
 DB 350 GCAGCATGCGCGGTATGAAGAAGGCTTTCGGGTTTGAAGCTTTCACTCGAGAG 409  
 QY 420 AAAAGGTAACGTAATTAATCTGACACCATGACGATGACAGAAAGACCGGCTAA 479  
 DB 410 AAGCGTTAGGTTAATTAATCTGACGATGATGATTAACCGGAGAAAGACCGGCTAA 469  
 QY 480 CTACGTCCAGCAGCGCGGTAATACGTAGGGTCAAGGCTTAATCGGAATTAATCGGGC 539  
 DB 470 CTCGGTCCAGCAGCGCGGTAATACGTAGGGTCAAGGCTTAATCGGAATTAATCGGGC 529  
 QY 540 TAAAGGTGCGAGCGCGCTTGAAGTCAATGTAAGTCAATCCCGGGCTTAACCTGGAA 599  
 DB 530 TAAAGGCGACGAGCGCGCTTGAAGTCAATGTAAGTCAATCCCGGGCTTAACCTGGAA 589  
 QY 600 TTGCGTTTGAATCTAACAAAGTGTGACAGAGGAGGATGGAATTCATGTGTAGCA 659  
 DB 590 CTGCAATTCGAACTGCGAGGCTAGAGCTTGTAGAGGGGGGTAGAAATTCAGGTTAGCG 649  
 QY 660 GTGAATGCGTAGAGATATGAAGAACATGATGCGAGAGCAGCTCTGGGTTAACAC 719  
 DB 650 GTGAATGCGTAGAGATCTGAAGAAATACGAGTGGCGAAGCGGCTCTGGGTTAACAC 709  
 QY 720 TGAAGCTATGCAAGAAAGGTGGGAGCAACAGATTAGATTAACCTGGTATGCCAGC 779  
 DB 710 TGAAGCTATGCAAGAAAGGTGGGAGCAACAGATTAGATTAACCTGGTATGCCAGC 769  
 QY 780 CTTAAAGCATGTAATCTAGTTGTTGGGCTTATTTAGGCTTGG-TAAAGAACTTAACCGT 838  
 DB 770 CTTAAAGCATGTAATCTAGTTGTTGGGCTTATTTAGGCTTGG-TAAAGAACTTAACCGT 829  
 QY 839 GAAAGTGAACGCTTGGGAGTACGCTGCAAGATTAAATCTAAAGAAATTGACGGGAG 898  
 DB 830 TAAAGTGAACGCTTGGGAGTACGCTGCAAGATTAAATCTAAAGAAATTGACGGGAG 889  
 QY 899 CCGGCAAGAGGCTGATTAATGATGATTAATTCATGCAAGGCAAGAAACCTTAACCTAAC 958  
 DB 890 CCGGCAAGAGGCTGATTAATGATGATTAATTCATGCAAGGCAAGAAACCTTAACCTAAC 949  
 QY 959 TTGAATGTAAGCAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAAGCTTAACAGAGTG 1017  
 DB 950 TTGAATGTAAGCAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAAGCTTAACAGAGTG 1009  
 QY 1018 CTGCAATGCTGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077  
 DB 1010 CTGCAATGCTGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1069  
 QY 1078 ACCCTTGTCTTAATTTGCCATC-ATTGGTTGGGCACTTTAATGAGCTGCGGTGCA 1136  
 DB 1070 ACCCTTGTCTTATTTGGTCCAGCGGTTAGGCGGGAACCTCAAGAGAGCTGCAAGTAA 1129  
 QY 1137 ACCGAGAGAGGTGGGATGACGTAAGCTCTTCAAGGCTTATGAGGTTAGGGCTTCAAC 1196  
 DB 1130 ACTGAGAGAGGTGGGATGACGTAAGCTCTTCAAGGCTTATGAGGTTAGGGCTTCAAC 1189  
 QY 1197 GTAATCAATGCGCGGTACAGAGGTTGCCAACCCGCGAGAGGAGCTTAATCTCAGAA 1256  
 DB 1190 GTGCTAATGCGCGGTACAGAGGTTGCCAACCCGCGAGAGGAGCTTAATCTCAGAA 1249  
 QY 1257 CCGGTCTATTCGGGATCGGAGTCTGCACTTCGATCCGTTGAAGTCCGAATCCGTTAGTA 1316  
 DB 1250 TGGGTCTATTCGGGATCGGAGTCTGCACTTCGATCCGTTGAAGTCCGAATCCGTTAGTA 1309  
 QY 1317 TCGGAGTACAGATGTCGCGGTGAATAGTTCCCGGGCTTGTGACACACCGCGCTCA 1374  
 DB 1310 TCGGAGTACAGATGTCGCGGTGAATAGTTCCCGGGCTTGTGACACACCGCGCTCA 1367

RESULT 64  
 US-10-808-807-16  
 ; Sequence 16, Application US/10808807  
 ; Patent No. 692928

GENERAL INFORMATION:  
 ; APPLICANT: E.I. du Pont de Nemours and Co., Inc.  
 ; APPLICANT: Cheng, Qiong  
 ; APPLICANT: Tao, Luan  
 ; APPLICANT: Sedkova, Natalia  
 ; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS  
 ; FILE REFERENCE: CL2365 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/808,807  
 ; PRIOR FILING DATE: 2003-06-12  
 ; PRIOR APPLICATION NUMBER: US 60/527,083  
 ; PRIOR FILING DATE: 2003-12-03  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: Patentin version 3.2  
 ; SEQ ID NO 16  
 ; LENGTH: 1324  
 ; TYPE: DNA  
 ; ORGANISM: Pantoea agglomerans strain DC404  
 US-10-808-807-16

Query Match 60.1%; Score 875.4; DB 3; Length 1324;  
 Best Local Similarity 81.9%; Pred. No. 2.4e-292;  
 Matches 1082; Conservative 0; Mismatches 231; Indels 8; Gaps 6;

QY 6 AGCTGGGGGCGATGCTTTACATGCAAGTGCAGCGGACGAGAT--GCTTGCACT- 62  
 DB 4 AGCTGGGGGCGAGGCTTTACATGCAAGTGCAGCGGACGAGATGCTTGCTGCTT 63  
 QY 63 -GGTGGAGTGGCGGAGCGGTGATGATGATCGAACGATTCAGAAAGGGGGTAA 121  
 DB 64 CGGTAGAGAGTGGCGGAGCGGTGATGATGATCGAACGATTCAGAAAGGGGGTAA 123  
 QY 122 CGCATGAAAGATGCTGTAATACCGGATTAATCTTAAGAGAAAGCAGGGGATGAA 181  
 DB 124 CTACTGAAACGATGCTGTAATACCGGATTAATCTTAAGAGAAAGCAGGGGATGAA 183  
 QY 182 ACCCTGCTTTTGGAGCGGCGGATGCTGATTAAGTATGATGATGATGATGATGATGATG 241  
 DB 184 GCTCTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 243  
 QY 242 AAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301  
 DB 244 TAGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303  
 QY 302 CCGAGCTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361  
 DB 304 TCCAGACTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 363  
 QY 362 AGCAATGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421  
 DB 364 AGCAATGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 423  
 QY 422 AAGGTTACGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 481  
 DB 424 GGTGTTAAGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 483  
 QY 482 AGTGGCAGGAGCGCGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 541  
 DB 484 CCGTGGCAGGAGCGCGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 543  
 QY 542 AAGGTTACGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601  
 DB 544 AAGGTTACGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 603  
 QY 602 GCGTTGAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661  
 DB 604 GCAATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663  
 QY 662 GAAATGCTTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 721  
 DB 664 GAAATGCTTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 723

QY 722 ACCCTCATGACGAAACCGTGGGAGCAACAGATTAGTACCTGTGTGTCACGGCC 781  
DB 724 ACCCTCAGGTCGCAAAACCGTGGGAGCAACAGATTAGTACCTGTGTGTCACGGCC 783  
QY 782 TAAACGATGCACTGTTGTTGGGCTTATTGAGGCTTG- TAAACGAGTTAACGCGTGA 840  
DB 784 TAAACGATGCACTGTTGTTGGGCTTATTGAGGCTTG- TAAACGAGTTAACGCGTGA 843  
QY 841 AGTTGACCGCTGGGAGTACGCTCGCAGATTAACTCAAGAAATTGACGCGGACCC 900  
DB 844 AGTGCAGCCGCTGGGAGTACGCGCCGCAAGTTAACTCAAGAAATTGACGCGGACCC 903  
QY 901 GCAACAGCGGTGATTAATGATGCAATTAATGATGCAACGCGGAAACCTTACTTACTT 960  
DB 904 GCAACAGCGGTGATGATGATGATTAATGATGCAACGCGGAAACCTTACTTACTT 963  
QY 961 GACATGAGGCAATTTTCTAGAGATGATAGTG- CTTGGGGAACGCTAACACAGGCTT 1019  
DB 964 GACATGAGGCAATTTTCTAGAGATGATAGTG- CTTGGGGAACGCTAACACAGGCTT 1023  
QY 1020 GCATGCTGTGTCAGCTCGTGTGATGATGATGATGATGATGATGATGATGATGATG 1079  
DB 1024 GCATGCTGTGTCAGCTCGTGTGATGATGATGATGATGATGATGATGATGATGATG 1083  
QY 1080 CTTTGTATTATTTGTCATC-ATTGTTGGGCACTTTAATGAGACTGCGGTGACAAAC 1138  
DB 1084 CTTTGTATTATTTGTCATC-ATTGTTGGGCACTTTAATGAGACTGCGGTGATTAAC 1143  
QY 1139 CGAGGAAAGGT- GGGGATGACGTCAATCTTCAATGAGGCTTAAAGGCTTCAACAG 1197  
DB 1144 CGAGGAAAGGTGAGGAGTGAAGTCAATGAGGCTTAAAGGCTTCAACAG 1203  
QY 1198 TAAATCAATGAGGAGTGAAGGCTTCAACGCGGAGGAGGAGTAACTCAAGAAAC 1257  
DB 1204 TGCTACATGAGGCTTCAACGAGAGAACTGCGGAGGAGGAGTAACTCAAGAAAC 1263  
QY 1258 GCGTGTGATGTCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1317  
DB 1264 ACCTGTGATGTCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1323  
QY 1318 C 1318  
DB 1324 C 1324

RESULT 65  
US-08-114-695A-2  
Sequence 2, Application US/08114695A  
GENERAL INFORMATION:  
PATENT No. 5508193  
APPLICANT: Mandelbaum, Raphael T.  
TITLE OF INVENTION: DEGRADATION OF S-TRIAZINES IN SOIL AND  
NUMBER OF INVENTIONS: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCHWEGMAN, LUNDBERG & WOESSNER, P.A.  
STREET: 3500 IDS CENTER  
CITY: MINNEAPOLIS  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/114,695A  
FILING DATE: 31-AUG-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MUEITING, ANN M.

REGISTRATION NUMBER: 33,977  
REFERENCE/DOCKET NUMBER: 600,268US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-339-0331  
TELEFAX: 612-339-3061  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1473 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: rRNA  
ORIGINAL SOURCE:  
ORGANISM: Pseudomonas sp. Atrazine-Degrading Isolate  
US-08-114-695A-2

Query Match 59.8%; Score 871.4; DB 2; Length 1473;  
Best Local Similarity 64.0%; Pred. No. 6,2e-291;  
Matches 920; Conservative 223; Mismatches 280; Indels 14; Gaps 9;

QY 17 ATGCTTTACATCATGCAATGTCGAAACGCGACACGATGCTTGCATCTGTGGCGATGGC 76  
DB 42 AGGCUAAACACATGCAAGUCGAGCGGAGGAGG- GCUUGCUUCCCGGCUUUAAGCGCG 100  
QY 77 GACGGGTGATTAATGATGCAATGTCGAAACGATTCAG- AGAGGGGGTAAACGATCGAAGATG 135  
DB 101 GAUGGUGUGAUAUUGCUAAGUAUUGCUUAGUGGAGGAGCAACGUCUCCAAAGAGAG 160  
QY 136 TGCTAATACCGCATATCTTAAAGAGAGAAAGGAGGAGATCGAAAGACTTGTGCGCTTTTG 195  
DB 161 CGCUAAUACCGCACUACGUCUCCCGGAGAGAAUUGGGAGUUCUUGGACCUACAGCUUUC 220  
QY 196 GAGCGCGCATGTCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 255  
DB 221 SAUGAGCUAAGGCGGAGUUAUGUAUUGUGGAGUUAUGGCUACCUAGGCGAGUACCG 280  
QY 256 TACTTGTGTCGAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 315  
DB 281 UAAUCUGUCUGAAGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 340  
QY 316 GAGGAGCAGAGTGGGAGATTTGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 375  
DB 341 GAGGAGCAGAGTGGGAGATTTGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 400  
QY 376 AGTGAAGAGCCCTTCGGGTTGTAAGCTTTTCACTGAGAGAAAGGTTACGCTAA 435  
DB 401 UGUGAAGAGAGUUCUGAGUUAAGCAUUAAGUUGGAGAGGAGCAUUAAGUUA 460  
QY 436 TAAATCGTGAACCATGACGATGATGACAGAAAGACACCGGCTTAACTGTCGACGACGCC 495  
DB 461 UACCUUGCUUUAUGAGUUAAGCAUUAAGCAUUAAGCAUUAAGCAUUAAGCAUUAAGCA 520  
QY 496 GCGGTAAATAC- GTTAGGTGCAAGCTTAACTGGAATTAATCTGAGGAGTAAAGGAGTGGCAG 554  
DB 521 GCGGUUAUACUAGAGGUGUACGUGUUAACGGAUUAUACGUGGUGUUAAGCGCGGAG 580  
QY 555 CGGCTTGTAAATGTCAGATGTAATCCCGGCTTAACTGAGAAATTGCGTTGAAACTA 614  
DB 581 UGUGUUGUUAAGUAGUUAAGCAUUAAGCAUUAAGCAUUAAGCAUUAAGCAUUAAGCA 640  
QY 615 CAAAGTGAAGTGTGACAGAGGAGGAGTGAATTCATGTGTAGAGTGAATTCGTTGAG 674  
DB 641 CCUGACUAGAGAGCGUAGAGGUGUGGAGUUAUUCUGUUAAGGAGGAGUUAUUCUGUUA 700  
QY 675 ATATGAGAAACATCATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 734  
DB 701 AUNGGAAGAAACACAGUGGCGAAAGGAGCACCUUGCAUUAUUAUUAUUAUUAUUAUUA 760  
QY 735 AAAGCGTGGAGAGCAACAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 794  
DB 761 CAAGC- UGGGAGGAGCAACAGUUAAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 819

QY	795	CTAGTGTGTTGGGCGCTTATTAGCGTTGTAACGAAGCTAAACGGGTGAAGTTGACCGGCTGG	854
DB	820	CUAGCGGTGGGAUUCUUGAGAUCCUAGUGGCGCAGUAAGGCGAUAAGUCGACCGCCCTGG	879
QY	855	GGAGTACGGTGCACAATTAAACTCAAGAAAGATTGACGGGGGACCCGCAACAAGCGTGGAA	914
DB	880	GCGGACGCGCCCAAGAGUUAACCAUAUGU---CAGGGGGCCGCAACCGGUGGA	936
QY	915	TTATGTGATTAATTCGATGCAACGCGAAAAACTTATCCTAACCTTGAACATGTACGAAT	974
DB	937	GCAUUGUGUUAUAUUCGAANNAAACCGAAGAACCUUACCTUGGCTCUGACAUUGCCGGAU	996
QY	975	TTTCTTAGAG-ATAGATTAGTGGCTTGGGGAAACCGTTAACACAGGTGCTGATGGCTGTCTC	1033
DB	997	CTUGCAGAGACAUCCAGAGAYCCUUGGGGAUUGGAAACACAGGUGGCUAGUGGCTUGCGUC	1056
QY	1034	AG---CTGCTGTCGTGAATGTTGGGTTAAAGTCCCGCAACGAGCGCCAAACCTGTGCTTA	1099
DB	1057	AGHUCAGUGGUCGAGAGAUUGUGGUAUAUGCCGUAACGAGCGCAACCUUGUCUUA	1116
QY	1091	ATTGCGATCA--TTTGGTTGGGCACTTTTAAATAGACTGCGGCTGCAAAACCGGA--GGAAG	1144
DB	1117	GUUACCAACGCUUACAGUGGGCAUCUAAGAGACUCCGSGUGACAAACCGGAAG	1176
QY	1148	GTGGGATGACGTCAAGTCTCATGACCCTTAATGGTAAAGGCTTACACAGTAAATTCATG	1207
DB	1177	GUGGGGAUGACGUCUACAUCAUAGGCCCCUUGCGGCCAGGGGCUAACACUGUCUUAUUG	1236
QY	1208	GCGCGTACAGAGGGTTGGCAACCGCGAGGGGAGCTAATCTCAAAAAGCGGTGTGTGT	1267
DB	1237	GUCGGUACAGAGGGUUGCCAAACCCGAGUGGAGCUUAUCCCAAGAAACCGAUGGUAUG	1296
QY	1268	CCGAGTCCGAGTCTGCAACTCCGTAAGTGGGAATCGCTAGTATCGCGGATCAG	1327
DB	1297	CCGGAUCCCAAGUCUCCACUCCAGCUGCGUAAGUGCGAUAUCCUUAUUGCGUAUACAG	1356
QY	1328	CATGTCCGCGTGAATACGTTCCCGGAGTCTTTGTACACACCGCCCGTACACCATGGAAGT	1387
DB	1357	AUUGCAGCGUAAUAACGUUCCCGGGCCUUGUACACACGCGCCGUCACACCAUUGGGAUGU	1416
QY	1388	GGTTTACCAACGAAGCAGTGAATCTAACCTTAAGAGAGGGCGCTTGCAACGCTGAAGTT	1444
DB	1417	GGUUCGUCUCCAGAAUGUACUAGCUAACGCGAAGGGGGAACGUGUACCAAGGAUGUAU	1473
RESULT 66			
US-09-198-955A-13			
; Sequence 13, Application US/09198955A			
; Patent No. 6187580			
; GENERAL INFORMATION:			
; APPLICANT: Andersen, Lene N.			
; APPLICANT: Schuelein, Martin			
; APPLICANT: Lange, Niels E.			
; APPLICANT: Bjornvad, Mads E.			
; APPLICANT: Moller, Soren			
; APPLICANT: Glad, Sanne O. S.			
; APPLICANT: Kauppinen, Markus S.			
; APPLICANT: Schorrt, Kirk			
; APPLICANT: Kongsbak, Lars			
; TITLE OF INVENTION: No. 6187580e1 Pectate Lyasee			
; FILE REFERENCE: 5378.200-US			
; CURRENT APPLICATION NUMBER: US/09198,955A			
; CURRENT FILING DATE: 1998-11-24			
; PRIOR APPLICATION NUMBER: 1343/97			
; PRIOR FILING DATE: 1997-11-24			
; PRIOR APPLICATION NUMBER: 1344/97			
; PRIOR FILING DATE: 1997-11-24			
; PRIOR APPLICATION NUMBER: 60/067,249			
; PRIOR FILING DATE: 1997-12-02			
; PRIOR APPLICATION NUMBER: 60/067,240			
; PRIOR FILING DATE: 1997-12-02			
; PRIOR APPLICATION NUMBER: 09/073,684			
; PRIOR FILING DATE: 1998-05-06			

	Prior Application Number: 09/164,217	
	1	Prior Filing Date: 1998-11-02
	2	Number of SEQ ID NOS: 32
	3	Software: FASTSEQ for Windows Version 4.0
	4	SEQ ID NO 13
	5	Length: 1506
	6	Type: RNA
	7	Organism: Bacillus sp. 1534
	8	US-09-198-955A-13
Query Match	58.3%; Score 850; DB 3; Length 1506;	
Best Local Similarity	61.8%; Pred. No. 1,7e+283;	
Matches	909; Conservative 215; Mismatches 330; Indels 16; Gaps 6	
DB	4	GAACGCTGCGGCATGCTTTACACATGCAAGTCGAAACGCGACGACGATGCTTGCATCTG 63
	4	GAACGCTGCGGCATGCTTTACACATGCAAGTCGAAACGCGACGACGATGCTTGCATCTG 63
DB	4	GAACGCTGCGGCATGCTTTACACATGCAAGTCGAAACGCGACGACGATGCTTGCATCTG 63
DB	64	GTGCGAATGCGCGACGCGGTGATGATGATCGG-AACGATCCGAAAGAGGGGGTAC 122
DB	64	GAUUAUACCGCGCGACGCGGTGATGATGATCGG-AACGATCCGAAAGAGGGGGTAC 123
DB	123	GCATCGAAAGATGCTTAAATCCGAT-ATACTCAAGAGGAAAGAGGGGATGAAAG 181
DB	124	UCCGGAAACCGGUCUUAUACCGGAUACACUUGAACCCUCGUGUCGAAAGUUGAAAG 183
DB	182	A-----CCTTGCGCTTTGGAGCGCGCATGTGTGATTGATCTGATTGGTGGGTA 231
DB	184	AUGGCGUUCGUGCUUACUUAAGAGUGGGCCGCGCGCAUUAAGUGUGUAAGUA 243
DB	222	AAAGCTTACCAAGCGGACGATCAGTATGTTGGTCTGAGAGACGACACCACTGGGAC 291
DB	244	AUGGCUUACCAAGCGGACGATCAGTATGTTGGTCTGAGAGACGACACCACTGGGAC 303
DB	292	TGAGCAAGCGCCGACGATCTTACCGGAGGACGACAGTGGGGAATTTGGACAAATGGCGCA 351
DB	304	UGAGCACGCGCCGACGATCTTACCGGAGGACGACAGTGGGGAATTTGGACAAATGGCGCA 363
DB	352	AGCGTATCCAGCAATGCGCGGTGATGAGAAAGGCTTTCGGTGTAAAGCTCTTTCAG 411
DB	354	AGUCUGACGAGACCAAGCGCGGTGATGAGAAAGGCTTTCGGTGTAAAGCTCTTTCAG 423
DB	412	TCGAGAGAAAGGTTACGG-TAAATTAATCGTACCCATGACCGGTATGCAAGAAAGAC 470
DB	424	GAGGAAAGAAAGAAUUCGUGUAUAAAGCGGUAUCUUGACGUAACUACCAAGAAC 483
DB	471	ACCGGCTAACTACGTCGACGACGCGCGGTATGATGAGGTGCAAGGGTTAAATGGGAT 530
DB	484	CACGGCUAACGTCGACGACGCGCGGTATGATGAGGTGCAAGGGTTAAATGGGAT 543
DB	531	TACTGAGCGTAAAGGATGCGACGCGCGCTTGTAGTACAGATGTTGAAATCCCGGGCTTA 590
DB	544	UAUUGGCGUAAAGCGCGCGGTGATGAGGTGCAAGGGTTAAATGGGATCCCGGGCTTA 603
DB	591	ACCTGGGAATTCGTTGAAATCTACAAAGCTAGAGTGTGGCAGAGGAGGTGAAATTCGA 650
DB	604	ACCGGACGCGCAUUGGAAACUUGGAGGACUUGAGUACGAAAGAGAGUUGAUAUCCA 663
DB	651	TGTGTAGCAGTGAATTCGTTGAAATCTACAAAGCTAGAGTGTGGCAGAGGAGGTGAAATTCGA 710
DB	664	CGUGTAGCGGUAUUGCGUAGUUAUUGGAGGAAACACAGUUGGCGAAAGGCGACUCUCG 723
DB	711	GTTTAACTACGTCGACGTCGACGAGAAAGCGTGGGAGGAAACAGGATTAAGTACCTGGT 770
DB	724	GUCUUAUACGUCGUGAGGCGCGGAAAGUGGGAGGAAACAGGATTAAGTACCTGGT 783
DB	771	AGTCACGCGCTTAAAGCAGTGTCAATGTTGCGCTTAAAGG--CTTGTGACGA 828
DB	784	AGUCCACGCGCTTAAAGCAGTGTCAATGTTGCGCTTAAAGG--CTTGTGACGA 843
DB	829	GCTTAAAGCGGTGAATGACCGCTTGGGAGTACGCTGCGACAGATTTAAACTCAAGGAT 888
DB	844	GUAACAACATUAAAGCAGTGTGCGGAGGAGTACGCTGCGACAGATTTAAACTCAAGGAT 903

Qy	889	TGACGGGAGCCCGCACAAAGCGGTGATTAATCTGATTAACTGCATGCAACGCCAAAAACC	948
Dd	904	UGAACGGGAGCCCGCACAAAGCAGUGAGCAGUCUGUUUAUUGCAGAACAACGCCAAGAACC	963
Qy	949	TTACTTAACTCTTGACATGTAGCCGAATTTTCTAGAGATAGA-TTAGTGCTTCGGGAACGCT	1007
Dd	964	UUAACGAGUGUCUGAACUCUUVUGACACUCUGGAGACAGAGCUCUCCCCUCGGGGGAAA	1023
Qy	1008	AACACAGGTGCTGATGCTGTGCTCATGCTGTGTGCTGAGATGTGTGGAGTTAACTCCGC	1067
Dd	1024	GUGACAGUGUGUGAUUGUGUUCUACAGUCUGUGUGUGAGAUUGUGUGUUAUCCCCGC	1083
Qy	1068	AACGAGCGCAACCTTGTCTATAATTCGCATCTTGTGGTGGGACCTTAAATGAACTGC	1127
Dd	1084	AACGAGCGCAACCCUUAUUCUUAUUGCAGACAUUUAUGUGGGCACUCUUAAGUGUAGCUCG	1143
Qy	1128	CGGTGACAAACCGGAGGAAAGGTGGGAGATGACGTCAAGTCTCTCATGCGCCCTTAATGGATAGG	1187
Dd	1144	CGUGUACAAACCGGAGGAAAGUGGGGGAGCAGAGCUUCAAUACAUCAGCCCUUAUACCUGG	1203
Qy	1188	GCTTCACACCGTAATACAAATGGCGGTCACAGGGGTGGCAACCCGCGAGGGGAGCTAAT	1247
Dd	1204	GCUACACCGUGCUAACAUUGAUUGGUACAAGGGUGGAAAGCCGCGAGUGAAAGCCAAT	1263
Qy	1248	CTCAGAAAGCCGCTCGTAGTCGCGATCGGATCTTGCAACTCGACTCCGTGAAGTCGGAAT	1307
Dd	1264	CCCUAUAAAGCCAUUUCUACGUCUGGACAUUGGCAAGCUCGACUCCUGACUAGAAGCCGGAU	1323
Qy	1308	CGCTTAGTAATCGCGGATCAGCAGATCTCGCGGTGAATACGTTCCCGGGTCTTGATACACCG	1367
Dd	1324	UGCUAUGUAUUGCCGGAUCAUUGCCGCGGUAUAUCUUCCCCGGGUCUUUUGUACACCG	1383
Qy	1368	CCCGTCACACCATGGGAATGGGTTTACCAACAAGCAGTACTCTAACCTGAAGAGAGGCG	1427
Dd	1384	CCCGUACACCAACGAGAUUUUGUACAACCCAAAGCCGUGAGGUAAACUUUUGAGGCAAG	1443
Qy	1428	CTTGGCACCGGTGAGATTGATGATCGGGGTG	1457
Dd	1444	CCGCCUAAAGUGGACAAAUAGAUUUUGGGGUG	1473

RESULT 67  
US-09-694-531-13  
Sequence 13, Application US/09654531  
Patent No. 6368843  
GENERAL INFORMATION:  
APPLICANT: Andersen, Lene N.  
APPLICANT: Schuelein, Martin  
APPLICANT: Lange, Niels E.  
APPLICANT: Bjornvad, Mads E.  
APPLICANT: Moller, Soren  
APPLICANT: Glad, Sanne O. S.  
APPLICANT: Kauplunen, Markus S.  
APPLICANT: Schmorck, Kirk  
APPLICANT: Kongsbæk, Lars  
TITLE OR INVENTION: No. 6368843el Pectate Lyases  
FILE REFERENCE: 5378.200-US  
CURRENT APPLICATION NUMBER: US/09/694, 531  
CURRENT FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 09/198, 955  
PRIOR FILING DATE: 1998-11-24  
PRIOR APPLICATION NUMBER: 1343/97  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 1344/97  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/067, 249  
PRIOR FILING DATE: 1997-12-02  
PRIOR APPLICATION NUMBER: 60/067, 240  
PRIOR FILING DATE: 1997-12-02  
PRIOR APPLICATION NUMBER: 09/073, 684  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 09/184, 217

```

: PRIOR FILING DATE: 1998-11-02
:
: NUMBER OF SEQ ID NOS: 32
:
: SOFTWARE: FastSeq for Windows
:
: SEQ ID NO 13
:
: LENGTH: 1506
:
: TYPE: RNA
:
: ORGANISM: Bacillus sp. 1534
:
: US-09-694-531-13

```

Query Match	58.3%	Score 850;	DB 3;	Length 1506;
Best Local Similarity	61.8%	Pred. No. 1.7e-283;		
Matches 909;	Conservative 215;	Mismatches 330;	Indels 16;	Gaps 6

OY	4	GAAGCTGTCGGCGATGCTTTTACATATGCATAGCTCAACCGGACGACACGGATGCTTGATCTG	63
Db	4	GAAGCTGTCGGCGATGCTTTTACATATGCATAGCTCAACCGGACGACACGGATGCTTGATCTG	63
OY	64	GTGGCGAGTGGCGGACCGGCTGATGTAATGTCATCGG--AACGTATCCAGAAAGAGGGGGTAAAC	122
Db	64	GAUUTUUGCCGCGGACCGGUGAUGAACAACGGGGCCMACCTGCCCTUUDGACUUGGGAUAAAC	123
OY	123	GCATCGAAGAATGTCATATACCGCAT--ATATCTCTAAGAGAGAAAAGCAGGGGATCGAAG	181
Db	124	UCCGGGAAACCGGUGCTUAAUACCGAGUAAACCTUUGAACCTCCGUGUUGCAAGUUGAAAG	183
OY	182	A-----CCTTGGCTTTTGGAGCGGCGCATGTCTGATTAGCTTAGTGTGGGTG	231
Db	184	AUGGCGUUCGUGCUAUCUAUAAAGAUUGGGCCCGCGCGCAUUAACUAGUUGUAAAGUA	243
OY	232	AAAGCTTACCAAGCGACGATCACTAATGTGTCTGAGAGGAGAGACACGACACACTGGGAC	291
Db	244	AUGGCUUACCAAGGCAAGGACGUAUGCCUAGCCGACCUAGAGGGUGAUCGGCCACACUUGGAC	303
OY	292	TGAGACACGCGCCACAGACTCTTACGCGAGACGACAGTGGGGAATTTTGGACAAATGGCGCA	351
Db	304	UGAAGACACGCGCCACAGACUUCUACCGGAGACGACAGUAGGAUACUUCGCGAAUUGACGAA	363
OY	352	AGCCTGATCCAGCAATGCCCGCTGAGTGAAGAAAGGCTTCCTGGCTTGAAGGTCTTTACG	411
Db	364	AGUCUGACGAGGACCAACGCGCGUGUGUAGGAAAGGCTCCUGGUGUAAAGCUCUGUUCU	423
OY	412	TCGAGAAAGAAAGGTTACGG--TAAATATCTGTAACCCATGACGATATGACAGAAAGAC	470
Db	424	GAGGAAAGAACAAAGUATCCGUGUAAUAAACCGGUACCTUAGACGUAACCUACACCAAGAAC	483
OY	471	ACCGGCTAATTAACGTGCCAGGACCGCGCGTATATATGTAAGGGTGCAGCGTTAATCGAAT	530
Db	484	CACGGCTUAAACUACGUGCCAGGACCGCGCGUAUAACGUAAGUGGCAAGCGUUCUCCGGAAU	543
OY	531	TACTGGCGGTAAAGGCTGCGACGCGCGCTTGTATAGTACAGATGTGAATATCCCGGGCTTA	590
Db	544	UAUUDGGGUGUAAAGCGCGCGACGCGGCGUUCUAAAGUCUGAAGUGUAAACUCGCGGCTCA	603
OY	591	ACCTGGGAATTGCGTTTGAATACTAACAAGCTTGAATGTGTGGCAGAGGAGAGTGAATTCCA	650
Db	604	ACCCCGAAGCGGCANUGGAAACUGGGGAGCUUAGUACGAAAGAGAGUGGAAUUCACA	663
OY	651	TGTGTAGCAGTGAATATGCGTAGAGATATGAGAAACAATCGATGGCGGAAGGACGCTCCG	710
Db	664	CGUGUAGCGGUGAAAUUGCGUAGAUUUGGAGAGAAACACAGUGGCGCAAGCGCACUCUCUG	723
OY	711	GCTTAAACCTAGCGCTCATGACGAAAGCGGTGGGAGCAAAACAGAGATTAGTATACCTGTGT	770
Db	724	GUCUUAACUACGCGUGAGCGCCGAAAGCGUUGGGAGCAAAACAGAGUUAUAGUAAUCCUGU	783
OY	771	AGTCCAGCGCTTAAACGATGTCAACTGATGTTGGCCTTATTAGG--CTTGGTAAACGA	828
Db	784	AGUCCACCGCGUAAACGAUAGAGUGCUAAGUGUUAAGGUGUUCGAGUCCGUGAGUGCCGAA	843
OY	829	GCTTACGCGTGAAGTTGAACCGGCTGGGGAGTACCGGTCGCAATTTAAACTCAAGAGAA	888
Db	844	GUAAACACUUAAGCACUCCGCTUGGGAGUACGACCGCACAGGUTUAAACUCCAAGGAU	903









Db 1109 ACAGCCCAACCTTGAATCTTAATGCGACATTTGAGATTGGGCACTTAAGTGAATGCC 1168  
QY 1129 GGTGCAAAACCGAGAGAGGTGGGATGAGCTGAAGTCTCATGCGCTTATGGTGAAGG 1188  
Db 1169 GGTGCAAAACCGAGAGAGGTGGGATGAGCTGAAGTCTCATGCGCTTATGGTGAAGG 1228  
QY 1189 CTTCACACTGAATATCAATGCGCGCTACAGAGGGTTCGCAACCGCGAGGGGGAGCTAAATC 1248  
Db 1229 CTACACACGCTCTACATAGGGTGTACMAAGGGCAGCAACGCGCAGAGCGGAGCAATC 1288  
QY 1249 TCAGAAAGCGCGTCTAGTCCGAGTCGAGTCTGCAACTGCACTCGTGAAGTCGGAATC 1308  
Db 1289 CAGAAAGCCACTCTCAATTCGATTCGAGCTGCGACTGCACTCGCTGATGAAGCGGAAT 1348  
QY 1309 GCTAGTAATGCGCGATCAAGCATTCGCGGTGAATACGTTCCGGGTCTTGTACACACCGC 1368  
Db 1349 GCTAGTAATGCGCGATCAAGCATTCGCGGTGAATACGTTCCGGGTCTTGTACACACCGC 1408  
QY 1369 CCGTCAACCATGAGGAGTGGGTTCACCAAGACAGTAAGCTTAACCTTAAGAGAGGGCGC 1428  
Db 1409 CCGTCAACCATGAGGAGTGGGTTCACCAAGACAGTAAGCTTAACCTTGTGAGCGAGC 1468  
QY 1429 TTGCCACGCTGAGATTCATGACTGGGCTG 1457  
Db 1469 CGCCGAAGGTGGGACAGATGATGGGCTG 1497

RESULT 70  
US-09-187-946-2/c  
Sequence 2, Application US/09187946  
Patent No. 6255467  
GENERAL INFORMATION:  
APPLICANT: Lindner, Luther E.  
APPLICANT: Macphree, Kathleen  
TITLE OF INVENTION: Human Blood Bacterium  
FILE REFERENCE: D6026  
CURRENT APPLICATION NUMBER: US/09/187, 946  
CURRENT FILING DATE: 1998-11-02  
EARLIER APPLICATION NUMBER: US 60/064,472  
EARLIER FILING DATE: 1997-11-06  
NUMBER OF SEQ ID NOS: 20  
SEQ ID NO 2  
LENGTH: 1502  
TYPE: DNA  
ORGANISM: unknown  
FEATURE:  
OTHER INFORMATION: 58 16S rRNA sequence of a new human blood bacterium  
US-09-187-946-2

Query Match 58.0%; Score 845.2; DB 3; Length 1502;  
Best Local Similarity 77.3%; Pred. No. 7.6e-282;  
Matches 1125; Conservative 1; Mismatches 274; Indels 56; Gaps 6;  
QY 4 GAAAGCTGGGCGATGCTTACACATGCAAGTCGAACGCGACGCGATGCTTGAATCTG 63  
Db 1500 GAAAGCTGGGCGATGCTTACACATGCAAGTCGAACGCGA-----CTTTC 1455  
QY 64 GTGGCAGATGCGCAGCGGTGAATGCAATCGAAAGTATCCAGAAAGAGGGGGTAAAG 123  
Db 1454 GGGGTCAAGCGCGAGCGGTGAATGCAAGCGGTGAAGAGTGTCCGTTCCGGAATTAAC 1395  
QY 124 CATGAAAGATGTGCTTAATCCGATTAATCTTAAGAGAGAAAGCAGGGATGAAAGAC 183  
Db 1394 CTGGGAAACTAGGCTTAATCCGATTAATCCCTTAATGG-----GAAAGGT 1349  
QY 184 CTGGCGCTTTTGAAGGGCGCGATGCTGATTAAGTATGTTGGTGAAGGCTTAACAA 243  
Db 1348 TTAATGCGGAAGATGCGCGCTGCTGATTAAGTATGTTGGTGAAGGCTTAACAA 1289  
QY 244 GGGCAGATCATGATGTTGCTCTGAGAGAGCAGCAGCAGCACTGGGACTGGAGCAAGGCC 303  
Db 1288 GGGCAGATCATGATGTTGCTCTGAGAGAGATGATCAGCAGCACTGGGACTGAGACAGGCC 1229

QY 304 CAGACTCTTACGGGAGGAGCAGATGGGGAAATTTTGGACAAATGGGCGCAAGCTGATCCAG 363  
Db 1228 CAGACTCTTACGGGAGGAGCAGATGGGGAAATTTTGGACAAATGGGCGCAAGCTGATCCAG 1169  
QY 364 CAATGCGCGGTGATGAAAGAGCCCTTCGGGTGTGAAGCTCTTTCACTGAGAAAGAAA 423  
Db 1168 CCATGCGCGGTGATGATGAAGCCCTTAGGGTGTGAAGCTCTTTTATCCGGGAGGATA- 1110  
QY 424 GGTACCGTAATTAATCGTACCCATGACCGTATGACAGAAAGAGCAGCGCTAACTAC 483  
Db 1109 -----ATGACGGTACCGAGGAATTTAGCCCGGCTTAATCTTC 1074  
QY 484 GTGCCAGACCGCGGTAAATAGTAGGGGTGAAGGCTTAATCGAATTAATCTGGGCGTAA 543  
Db 1073 GTGCCAGACCGCGGTAAATAGTAGGGGTGAAGGCGCTTAATCGAATTAATCTGGGCGTAA 1014  
QY 544 GGGTGGCAGGCGCGCTTGTAAAGTGAAGTGAATCCCGGGCTTAACTGGGAAATTC 603  
Db 1013 GGGCGGTATGCGGGCTTTTAAGTCGGGGGTGAAGCCTGTGGCTCAACAGAAATGCG 954  
QY 604 GTTTGAAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 663  
Db 953 CTTCGATCTGGGAGCGCTGAGTATGTAAGGTTGGTGAAGTGAAGTGAAGTGAAGTGAAG 894  
QY 664 AATGCGTAGAGATATGGAAGAACATGATGCGGAAGCGAGCTCTGGGTAACTGAC 723  
Db 893 AATGCGTAGATATTTGCAAGAACACCGGTGGGAGCGGCAACTGACATTAATCTGAC 834  
QY 724 GCTCATGACGAAGAGGTGGGAGCAACAGAGTTAGATACCTGTGATGCCAGCCCTTA 783  
Db 833 GCTGAGGCGCGAAGAGGTGGGAGCAACAGAGTTAGATACCTGTGATGCCAGCCCTTA 774  
QY 784 AACGATGTAATGATGTTGGGCTTATAGGCTTGTGAAGAGCTTAACCGGTGAAGT 843  
Db 773 AACGATGTAATGCAAGTGTGGGCTTGTGACCGCAAGTGAAGCTTAACCGGTGAAGT 714  
QY 844 TGACCGCGCTGGGAGATGCGTGCAGAGTAAATCTCAAGAGATGACGGGAGCCGCA 903  
Db 713 ATTCGCGCTGGGAGATGCGTGCAGAGTAAATCTCAAGAGATGACGGGAGCCGCA 654  
QY 904 CAAGCGGTGATTAATGATTAATTCATGATGCAAGCGCAAAACCTTAACCTTGAAC 963  
Db 653 CAAGCGGTGATTAATGATTAATTCATGATGCAAGCGCAAAACCTTAACCTTGAAC 594  
QY 964 ATGACCGAATTTTCTAGAGA-TAGATTAAGTCTTGGGAAGCTTAACAGAGTCTGA 1022  
Db 593 ATGACCGTATTAACCAAGAGATTTGGGATCTCACTTCGATGGGCGGCAAGAGTCTGA 534  
QY 1023 TGGCTGTGTCAGCTGCTGCTGATGATGTTAGTGGGTTAAGTCCGCAACAGCGCAACCT 1082  
Db 533 TGGCTGTGTCAGCTGCTGCTGATGATGTTAGTGGGTTAAGTCCGCAACAGCGCAACCT 474  
QY 1083 TGTCAATTAATTCATCAATTTGTTGGGCACTTAAATGAGACTGCGGTGAACACG-G 1141  
Db 473 CGCTCTTAATGTCATCAATTTGTTGGGCACTTAAATGAGACTGCGGTGAACACG-G 414  
QY 1142 AGGAAGTGGGATGACGTCAGTCTCTCATGCGCTTAATGAGTGGGCTTCAACAGTAA 1201  
Db 413 AGGAAGTGGGATGACGTCAGTCTCTCATGCGCTTAAATGAGTGGGCTTCAACAGTAA 354  
QY 1202 ACAATGCGCGTACAGAGGTTGGCAACCGGAGAGGGAGCTTAATCTCAGAAAGCGCGT 1261  
Db 353 ACAATGCGCGTACAGAGGTTGGCAACCGGAGAGGGAGCTTAATCTCAGAAAGCGCGT 295  
QY 1262 CGTATGCGCGATGAGTGTGCACTGCACTCGTGAAGTGGAAATCGTATGATCGCG 1321  
Db 294 CTCAGTGTGATTTGCACTTGCACCTGAGTCAATGAAGCGGAAATCGTATTAATCGTG 235  
QY 1322 GATCAGCATGTGCGGTGAATACGTTCCGGGTCTTTGTAACAACCGCGCTGCAACCAAG 1381  
Db 234 GATCAGCATGTGCGGTGAATACGTTCCGGGTCTTTGTAACAACCGCGCTGCAACCAAG 175  
QY 1382 GGAAGTGGTTTACCAAGAGAGTATGCTTAACCGTAAGAGAGGCGCTTGCCAGCGTGA 1441

Mon May 8 11:07:21 2006

us-10-659-980a-2.rni

Page 69

D6  
174 GGACTTGGTCCTTAACCGAGCGGCGCTGCAGCAACCGCAGAAGGCAAGGCCAACCAGGTAAAG 115

OY  
1442 ATTCACTACTGGGGTG 1457

D8  
114 GTCAAGCACTGGGGTG 99

Search completed: May 5, 2006, 13:24:46  
Job time : 313 secs

**THIS PAGE LEFT BLANK**



```
;
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type A1 R7clone187 16S rDNA
US-10-659-948A-2

Query Match      100.0%; Score 1457; DB 7; Length 1457;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTGAAGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGAGCATCTTGCAAT 60
DB 1 ATTGAAGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGAGCATCTTGCAAT 60
QY 61 CTGGTGGCGAGTGGCGGACGGGTAGTAAATGCAATCGGAAGGTAAGAGGAGGAGT 120
DB 61 CTGGTGGCGAGTGGCGGACGGGTAGTAAATGCAATCGGAAGGTAAGAGGAGGAGT 120
QY 121 ACCGATCGAAAGATGTGCTAAATACCGCATTAATCTTAAGAGGAAAGCAGGGATGAAA 180
DB 121 ACCGATCGAAAGATGTGCTAAATACCGCATTAATCTTAAGAGGAAAGCAGGGATGAAA 180
QY 181 GACCTTGCGCTTTGGAGCGGCGGATGCTGATTAAGCTAGTGGGAGTAAAGGCTTAC 240
DB 181 GACCTTGCGCTTTGGAGCGGCGGATGCTGATTAAGCTAGTGGGAGTAAAGGCTTAC 240
QY 241 CAAGGCGAGCATAGTATGTTGCTGAGAGAGCAGCAGCAGCAGCTGGAGCTGAGACAG 300
DB 241 CAAGGCGAGCATAGTATGTTGCTGAGAGAGCAGCAGCAGCAGCTGGAGCTGAGACAG 300
QY 301 GCCCAACTCTTACCGGAGGAGCAGATGGGGAATTTTGGCAATGGGCGGAAGCCTGATC 360
DB 301 GCCCAACTCTTACCGGAGGAGCAGATGGGGAATTTTGGCAATGGGCGGAAGCCTGATC 360
QY 361 CAGCAATGCGCGTGAAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGAAGAGA 420
DB 361 CAGCAATGCGCGTGAAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGAAGAGA 420
QY 421 AAAGGTTACGGTAAATATATGTAACCCATGACGGTATCGACAGAAAGACCGGCTTAC 480
DB 421 AAAGGTTACGGTAAATATATGTAACCCATGACGGTATCGACAGAAAGACCGGCTTAC 480
QY 481 TACGTGCGCAGACCGCGGTAAATACGTAAGGTCGAAAGCTTAACTGAGGCGT 540
DB 481 TACGTGCGCAGACCGCGGTAAATACGTAAGGTCGAAAGCTTAACTGAGGCGT 540
QY 541 AAAGGAGCGCAGGCGGCTTGTAAAGTCAGATGTAAGTCCCGGCTTAACTGAGGAAT 600
DB 541 AAAGGAGCGCAGGCGGCTTGTAAAGTCAGATGTAAGTCCCGGCTTAACTGAGGAAT 600
QY 601 TGGGTTTGAATCTAAAGCTAGAGTGTGCAAGGAGGTGAAATCCCATGTGTAGCAG 660
DB 601 TGGGTTTGAATCTAAAGCTAGAGTGTGCAAGGAGGTGAAATCCCATGTGTAGCAG 660
QY 661 TGAATATGCGTAAGATATGGAAGAACTCATGATGCGCAAGGCAAGCTTCCTGGTTAACT 720
DB 661 TGAATATGCGTAAGATATGGAAGAACTCATGATGCGCAAGGCAAGCTTCCTGGTTAACT 720
QY 721 GACGCTCATGCAAGAAAGCTGGGAGCAAAAGGATTAATATCCCTGATGTCAGCGCC 780
DB 721 GACGCTCATGCAAGAAAGCTGGGAGCAAAAGGATTAATATCCCTGATGTCAGCGCC 780
QY 781 CTAAACGATGTCAACTAGTGTGGGCTTATTAAGCTTGTAAAGAGTAAACGCGTGA 840
DB 781 CTAAACGATGTCAACTAGTGTGGGCTTATTAAGCTTGTAAAGAGTAAACGCGTGA 840
QY 841 AGTTGACCGCTGGGAGATCGGTGCGAAGATTAAACTAAAGAAATTGAACGGGACCC 900
DB 841 AGTTGACCGCTGGGAGATCGGTGCGAAGATTAAACTAAAGAAATTGAACGGGACCC 900
QY 901 GCACAAGCGGTGATTAATGATTAATGATGCAACGGGAAACCTTAACCTTACCTT 960
DB 901 GCACAAGCGGTGATTAATGATTAATGATGCAACGGGAAACCTTAACCTTACCTT 960
```

```
DB 901 GCACAAGCGGTGATTAATGATTAATGATGCAACGGGAAACCTTAACCTTACCTT 960
QY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGTCTGGGAAAGCTAAACAGGTGCTG 1020
DB 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGTCTGGGAAAGCTAAACAGGTGCTG 1020
QY 1021 CATGGCTGTCTCAGCTGCTGTCTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCAAC 1080
DB 1021 CATGGCTGTCTCAGCTGCTGTCTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGCAAC 1080
QY 1081 CTGTTCATTAATTCCTCATTTGGTGGGCACTTAAATGAGACTGCGGCTGACAAACCG 1140
DB 1081 CTGTTCATTAATTCCTCATTTGGTGGGCACTTAAATGAGACTGCGGCTGACAAACCG 1140
QY 1141 GAGGAAGTGGGAGATGACCTCAAGTCCCTTAATGAGTGGGCTTCAACGTTAA 1200
DB 1141 GAGGAAGTGGGAGATGACCTCAAGTCCCTTAATGAGTGGGCTTCAACGTTAA 1200
QY 1201 TACAAATGGCGCTTACAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAACCGCG 1260
DB 1201 TACAAATGGCGCTTACAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAACCGCG 1260
QY 1261 TCGTATGCTCGGATTCGGAATCTGCAACTGCACTCCGTGAAGTGGATCTGTATTCGC 1320
DB 1261 TCGTATGCTCGGATTCGGAATCTGCAACTGCACTCCGTGAAGTGGATCTGTATTCGC 1320
QY 1321 GGATGACGATGTCGGGTGAATACGTTCCGGGCTCTGTAACAACCGCGCTGACACCAT 1380
DB 1321 GGATGACGATGTCGGGTGAATACGTTCCGGGCTCTGTAACAACCGCGCTGACACCAT 1380
QY 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTAATCCGTAAAGAGGCGCTTGGCAACGTTGA 1440
DB 1381 GGGAGTGGGTTTCAACAGAGAGGATGCTAATCCGTAAAGAGGCGCTTGGCAACGTTGA 1440
QY 1441 GATTATGACTGGGGTG 1457
DB 1441 GATTATGACTGGGGTG 1457

RESULT 2
US-10-659-980A-2
; Sequence 2, Application US/10659980A
; Publication No. US20040106133A1
; GENERAL INFORMATION:
; APPLICANT: Hovavac, Timothy A
; TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284/781
; CURRENT APPLICATION NUMBER: US/10/659,980A
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type A1 R7clone187 16S rDNA
US-10-659-980A-2

Query Match      100.0%; Score 1457; DB 7; Length 1457;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTGAAGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGAGCATCTTGCAAT 60
DB 1 ATTGAAGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGAGCATCTTGCAAT 60
```



OY	241	CAGGCCACGANTCACTAGTTGGTCTGTAGAAGACAACCAACCGACAATGGAGATTGAGAACG	300
Db	241	CAAGCCGACGATCAGTAGTTGGTCTGTAGAAGACAACGACCACACTGGGAATGTGACACGC	300
OY	301	GCCCCAGATCCCTCAACGGAGAGCGACGATGGGGAAATTTTGACAAATGGGGGCAAGCCGTATC	360
Db	301	GCCCAGACTCTTAACGGAGGACGACAGTGGGAATTTTGGACAAATGGGGCGCAAGCTGTATC	360
OY	361	CAGCAATGCCCGGTGAGTGAAGAAAGGCCCTTCGGGTTGTAAAGCTCTTTCAGTGCAGAAGA	420
Db	361	CAGCAATGCCCGGTGAGTGAAGAAAGGCCCTTCGGGTTGTAAAGCTCTTTCAGTGCAGAAGA	420
OY	421	AAAAGTTACGGTAAATAATTCGTGACCCATGACGGTATGACAGAAAGAACCGGCTAAC	480
Db	421	AAAAGTTACGGTAAATAATTCGTGACCCATGACGGTATGACAGAAAGAACCGGCTAAC	480
OY	481	TACGTGCCAGACGCCCGGTAAATACGTAGGGGTCAAAGGTTAAATCGAAATTACTGGGGGT	540
Db	481	TACGTGCCAGACGCCCGGTAAATACGTAGGGGTCAAAGGTTAAATCGAAATTACTGGGGGT	540
OY	541	AAAAGGTCCGACGAGCGGCTTGTAAAGTGCAGATGTGAAATCCCGGGCTTAACTGGGAAT	600
Db	541	AAAAGGTCCGACGAGCGGCTTGTAAAGTGCAGATGTGAAATCCCGGGCTTAACTGGGAAT	600
OY	601	TGCGTTTAAACTACAAAGCTAAGATGTGGCAGAGGAGGTGAAATTCATGTGTAGACG	660
Db	601	TGCGTTTAAACTACAAAGCTAAGATGTGGCAGAGGAGGTGAAATTCATGTGTAGACG	660
OY	661	TGAAATGGGTAGAGTATGAAAGAACATCGAGTGGGAAGGACGCTCTGGGTTTAACT	720
Db	661	TGAAATGGGTAGAGTATGAAAGAACATCGAGTGGGAAGGACGCTCTGGGTTTAACT	720
OY	721	GACGCTCATGACAGAAACGTTGGGAGCAAAACAGATTATGATTACCTGTGATGCACGCC	780
Db	721	GACGCTCATGACAGAAACGTTGGGAGCAAAACAGATTATGATTACCTGTGATGCACGCC	780
OY	781	CTAAACGATGTCAACTAGTTGTTGGGCTTATTAGGCTTGTGTACGAAGCTAACGCGTGA	840
Db	781	CTAAACGATGTCAACTAGTTGTTGGGCTTATTAGGCTTGTGTACGAAGCTAACGCGTGA	840
OY	841	AGTTGACCGCCCTGGGGAGTAGACGCTGCGAAAGTTAAATCTCAAAGGAATTTGACGGGGACCC	900
Db	841	AGTTGACCGCCCTGGGGAGTAGACGCTGCGAAAGTTAAATCTCAAAGGAATTTGACGGGGACCC	900
OY	901	GCACAAGCGGTGATATGTGATTAATTCGATGCACGCGCAAAAACTTACTACTACCTT	960
Db	901	GCACAAGCGGTGATATGTGATTAATTCGATGCACGCGCAAAAACTTACTACTACCTT	960
OY	961	GACATGTAGCGAATTTTCTAGAGATAGATTAGTGTCTGGGAACGCTAACACAGTGTG	1020
Db	961	GACATGTAGCGAATTTTCTAGAGATAGATTAGTGTCTGGGAACGCTAACACAGTGTG	1020
OY	1021	CATGCGTGTGCTCAGCTGTGTGCTGTGAGATGTGGGTTAATGCTCCGCAACGAGGCAACC	1080
Db	1021	CATGCGTGTGCTCAGCTGTGTGCTGTGAGATGTGGGTTAATGCTCCGCAACGAGGCAACC	1080
OY	1081	CTTGTCATTAATTTGCAATCTTTGGTGGGCACTTAATGGAAGTCCGCGTGACAAACCG	1140
Db	1081	CTTGTCATTAATTTGCAATCTTTGGTGGGCACTTAATGGAAGTCCGCGTGACAAACCG	1140
OY	1141	GAGGAAGGTGGGAGTAGCGTCAAGTCTCATGGCCCTTAATGGGTATGGGCTTCAACGTTAA	1200
Db	1141	GAGGAAGGTGGGAGTAGCGTCAAGTCTCATGGCCCTTAATGGGTATGGGCTTCAACGTTAA	1200
OY	1201	TACAAATGGCGCGTACAGAGGGTTGCAACCCGCGAGGAGGAGACTAATCTCAGAAAGCGCG	1260
Db	1201	TACAAATGGCGCGTACAGAGGGTTGCAACCCGCGAGGAGGAGACTAATCTCAGAAAGCGCG	1260
OY	1261	TGTTAGTCCGATCCGAGTCTGCAACTGCACTTCGTTGAAGTCCGAATGCGTAGTAATCG	1320
Db	1261	TGTTAGTCCGATCCGAGTCTGCAACTGCACTTCGTTGAAGTCCGAATGCGTAGTAATCG	1320

QY	1321	GGATTCAGCATGTCCGGCTGGAATACGTTCCCGGGCTTTGTATACACGCCGCCGTACACCAT	1380
DB	1321	GGATTCAGCATGTCCGGCTGGAATACGTTCCCGGGCTTTGTATACACGCCGCCGTACACCAT	1380
QY	1381	GGGATGGGGTTTACACGAAGCAGTAGTCTAACCGTTAAGAGAGGCCCTTGCCACGGTGA	1440
DB	1381	GGGATGGGGTTTACACGAAGCAGTAGTCTAACCGTTAAGAGAGGCCCTTGCCACGGTGA	1440
QY	1441	GATTCATGACTGGGGTG	1457
DB	1441	GATTCATGACTGGGGTG	1457
RESULT 4			
US-10-659-948A-1			
; Sequence 1, Application US/10659948A			
; Publication No. US20040101946A1			
GENERAL INFORMATION:			
APPLICANT: Hoevec, Timothy A			
TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria			
FILE REFERENCE: 81289-294309			
CURRENT APPLICATION NUMBER: US/10/659, 948A			
CURRENT FILING DATE: 2003-09-10			
PRIOR APPLICATION NUMBER: US 09/573,684			
PRIOR FILING DATE: 2000-05-19			
PRIOR APPLICATION NUMBER: US 60/386,217			
PRIOR FILING DATE: 2002-09-19			
PRIOR APPLICATION NUMBER: US 60/386,218			
PRIOR FILING DATE: 2002-09-19			
PRIOR APPLICATION NUMBER: US 60/386,219			
PRIOR FILING DATE: 2002-09-19			
NUMBER OF SEQ ID NOS: 23			
SOFTWARE: PatentIn version 3.2			
SEQ ID NO 1			
LENGTH: 1457			
TYPE: DNA			
ORGANISM: Unknown			
FEATURE:			
OTHER INFORMATION: AOB Type A R7C1one140 16S rDNA			
US-10-659-948A-1			
Query Match 99.7%; Score 1452.2; DB 7; Length 1457;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	1	ATTGAACCTCGCGGCACTCTTTACATGGAATCGAACGGCGAGCCGATGCTTGAT	60
DB	1	ATTGAACCTCGCGGCACTCTTTACATGGAATCGAACGGCGAGCCGATGCTTGAT	60
QY	61	CTGGTGGCGAATGGCGGAGCGGGTAGTAGTATCATCGNACCTATCCAGAAAGGGGGGTA	120
DB	61	CTGGTGGCGAATGGCGGAGCGGGTAGTAGTATCATCGNACCTATCCAGAAAGGGGGGTA	120
QY	121	ACGCATCGAAGATGTGCTAATACCGCATATATCTTAAAGAGAAAGCAGGGGATCGAA	180
DB	121	ACGCATCGAAGATGTGCTAATACCGCATATATCTTAAAGAGAAAGCAGGGGATCGAA	180
QY	181	GACCTTGGGCTTTTGGAGCGGGCCGATGTCTGATTAAGTAGTAGTTGGGGGGTAAAGGCTTAC	240
DB	181	GACCTTGGGCTTTTGGAGCGGGCCGATGTCTGATTAAGTAGTAGTTGGGGGGTAAAGGCTTAC	240
QY	241	CAAGCGCAGATCAGTAGTATGTTGGTCTGAGAGAGACGACCGCACACTGGGACTGAGACAG	300
DB	241	CAAGCGCAGATCAGTAGTATGTTGGTCTGAGAGAGACGACCGCACACTGGGACTGAGACAG	300
QY	301	GCCCAAGCTCTTACCGGAGGCGACAGTAGTGGGAATTTTGGACAATGGGGCGCAAGCTTATC	360
DB	301	GCCCAAGCTCTTACCGGAGGCGACAGTAGTGGGAATTTTGGACAATGGGGCGCAAGCTTATC	360
QY	361	CAGCAATCCCGGTGAGTGAAGAAAGGGCTTGGGGGTGTAAGCTCTTTCACTCGAAGA	420
DB	361	CAGCAATCCCGGTGAGTGAAGAAAGGGCTTGGGGGTGTAAGCTCTTTCACTCGAAGA	420



QY	421	AAAGGTTACGGTAAATTAATCGTGA	CCCATGACGGTATCTGACAGAAAGAAAGACCGGCTAAC	480
Db	421	AAAGGTTACGGTAAATTAATCGTGA	CCCATGACGGTATCTGACAGAAAGAAAGACCGGCTAAC	480
QY	481	TACGTGCACAGACCGCGGTAA	TACGTAGGGTGCAGCGTTAATCGGAATTTACTGAGCGT	540
Db	481	TACGTGCACAGACCGCGGTAA	TACGTAGGGTGCAGCGTTAATCGGAATTTACTGAGCGT	540
QY	541	AAAGGGTGCAGACCGGCTTGT	AAAGTCAGATGTGAAATCCCGGGCTTAACTGCGGAAT	600
Db	541	AAAGGGTGCAGACCGGCTTGT	AAAGTCAGATGTGAAATCCCGGGCTTAACTGCGGAAT	600
QY	601	TGCGTTTGAAACTACAAAGCT	AGTAGGTGGCAGAGGAGGTGGAAATTCATGTGTACAG	660
Db	601	TGCGTTTGAAACTACAAAGCT	AGTAGGTGGCAGAGGAGGTGGAAATTCATGTGTACAG	660
QY	661	TGAAATGGGTAGAGTATG	AGAAACATCGATGGCGAAGGACGCTCTGGGGTTAACT	720
Db	661	TGAAATGGGTAGAGTATG	AGAAACATCGATGGCGAAGGACGCTCTGGGGTTAACT	720
QY	721	GACGCTCATGACAGAAACG	TGGGAGCAACAGATTAAGTATACCTGTAGTCACGCC	780
Db	721	GACGCTCATGACAGAAACG	TGGGAGCAACAGATTAAGTATACCTGTAGTCACGCC	780
QY	781	CTAAACGATGTCAACTAG	TGTGTGGGCTTATTAAGCTTGGTAAACGAAGCTTAACGCGTGA	840
Db	781	CTAAACGATGTCAACTAG	TGTGTGGGCTTATTAAGCTTGGTAAACGAAGCTTAACGCGTGA	840
QY	841	AGTTAACCGGCTTGGGAG	TACCGTTCGCAAGATTAAACTCAAAAGAAATTGACGGGGAACCC	900
Db	841	AGTTAACCGGCTTGGGAG	TACCGTTCGCAAGATTAAACTCAAAAGAAATTGACGGGGAACCC	900
QY	901	GCACAAGCGGTAGTTAT	TGTGGAATTAAATTCGATGCACGCGCAAAAAACCTTACTACCTT	960
Db	901	GCACAAGCGGTAGTTAT	TGTGGAATTAAATTCGATGCACGCGCAAAAAACCTTACTACCTT	960
QY	961	GACATGTAGCGAATTTT	TAGAGATGAGTTAGTGTCTGGGAAACGCTTAACAACAGTGTCTG	1020
Db	961	GACATGTAGCGAATTTT	TAGAGATGAGTTAGTGTCTGGGAAACGCTTAACAACAGTGTCTG	1020
QY	1021	CATGGCTGTCTGACGCT	GTGTGCTGATGATGTTGGGTTAAATGCCCGCAACAGAGCGCAACC	1080
Db	1021	CATGGCTGTCTGACGCT	GTGTGCTGATGATGTTGGGTTAAATGACACTGCCGCGTGAACAACG	1080
QY	1081	CTTGTGATTAATTTGG	CATCATTTGGTGGGCACTTAAATGAGACTGCCGCGTGAACAACG	1140
Db	1081	CTTGTGATTAATTTGG	CATCATTTGGTGGGCACTTAAATGAGACTGCCGCGTGAACAACG	1140
QY	1141	GAGGAAGGTGGGAGTAG	CGTCAAGTCTTCATAGGCCCTTAATGGGTAGGGCTTCAACGTAA	1200
Db	1141	GAGGAAGGTGGGAGTAG	CGTCAAGTCTTCATAGGCCCTTAATGGGTAGGGCTTCAACGTAA	1200
QY	1201	TACAAATGCGCGGTAC	AGAGGGTTGCCAACCCGCGAGGGGAGACTTAATCAGAAAGCGCG	1260
Db	1201	TACAAATGCGCGGTAC	AGAGGGTTGCCAACCCGCGAGGGGAGACTTAATCAGAAAGCGCG	1260
QY	1261	TGCTAGTCCGGATCG	AGTCTCGAATCTCGAATCCGGAAGTGGAAATGCTAGTAATGCG	1320
Db	1261	TGCTAGTCCGGATCG	AGTCTCGAATCTCGAATCCGGAAGTGGAAATGCTAGTAATGCG	1320
QY	1321	GGATCAGCAATGTGCG	CGGTAAACGTTCCGGGGCTTGTGAACAACCGGCCCTGACACCAT	1380
Db	1321	GGATCAGCAATGTGCG	CGGTAAACGTTCCGGGGCTTGTGAACAACCGGCCCTGACACCAT	1380
QY	1381	GGGAGTGGGTTTCA	CCAGAGCAGTATGCTTAAACCGTAAAGAGGGGCGCTTGCCACGGTGA	1440
Db	1381	GGGAGTGGGTTTCA	CCAGAGCAGTATGCTTAAACCGTAAAGAGGGGCGCTTGCCACGGTGA	1440
QY	1441	GATTCAATGACTGGGGT	1457	
Db	1441	GATTCAATGACTGGGGT	1457	

```

RESULT 5
US-10-659-980A-1
; Sequence 1, Application US/10659980A
; Publication No. US20040106133A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneec, Timothy A
; TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284781
; CURRENT APPLICATION NUMBER: US/10/659, 980A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573, 664
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386, 217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386, 218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386, 219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type A R7clone140 16S rDNA
US-10-659-980A-1

```

Query Match	99.7%	Score 1452.2	DB 7	Length 1457
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1454	Conservative	0	Mismatches 3	Indels 0
				Gaps 0
Qy	1	ATTGAACCTGGCCGATGCTTTACATCAATGAAAGTCGAAACGGCAGCACGATGCTTGCAT	60	
Db	1	ATTGAACCTGGCCGATGCTTTACATCAATGAAAGTCGAAACGGCAGCACGATGCTTGCAT	60	
Qy	61	CTGTGTGGCGAGTGGCGGACGGGGTAGTAATGCATCGAAAGTATTCAGAAAGAGGGGGTA	120	
Db	61	CTGTGTGGCGAGTGGCGGACGGGGTAGTAATGCATCGAAAGTATTCAGAAAGAGGGGGTA	120	
Qy	121	ACGCATCGAAAGATGTGCTAATACCGCATATACTCTAAGAGAGAAAGCAGGGGATCGAAA	180	
Db	121	ACGCATCGAAAGATGTGCTAATACCGCATATACTCTAAGAGAGAAAGCAGGGGATCGAAA	180	
Qy	181	GACTTGGCGCTTTTGAAGCGGCGCATGTCTGATTAGTAGTTGTGGGCTTAAAGGCTTAC	240	
Db	181	GACTTGGCGCTTTTGAAGCGGCGCATGTCTGATTAGTAGTTGTGGGCTTAAAGGCTTAC	240	
Qy	241	CAAGCGACGATCACTAGTTGGTCTGAGAGAGACGACCGACCACTTGGGACTGAGACACG	300	
Db	241	CAAGCGACGATCACTAGTTGGTCTGAGAGAGACGACCGACCACTTGGGACTGAGACACG	300	
Qy	301	GCCCAGACTCTTACGSGAGGACGACGATGGGGATTTTGGACATGGGGCGCAAGCCTGATC	360	
Db	301	GCCCAGACTCTTACGSGAGGACGACGATGGGGATTTTGGACATGGGGCGCAAGCCTGATC	360	
Qy	361	CAGCAATGCGCGGTGATGTAAGAGGCTTCCGGTTGTAAGCTCTTTCACTCGAGAAAGA	420	
Db	361	CAGCAATGCGCGGTGATGTAAGAGGCTTCCGGTTGTAAGCTCTTTCACTCGAGAAAGA	420	
Qy	421	AAAGGTTAGGTAATTAATCGTACCCATGACGATGACGAGAAAGCACCGGCTAAC	480	
Db	421	AAAGGTTAGGTAATTAATCGTACCCATGACGATGACGAGAAAGCACCGGCTAAC	480	
Qy	481	TACGTGCCAGACCGCGGCTAATACGTAGAGGTGCAAGCGTTAATCGAAATTACTGGCGCT	540	
Db	481	TACGTGCCAGACCGCGGCTAATACGTAGAGGTGCAAGCGTTAATCGAAATTACTGGCGCT	540	
Qy	541	AAAGGTTAGGTAATTAATCGTACCCATGACGATGTAAGTAAATCCCGGGCTTAACTGGGAAT	600	
Db	541	AAAGGTTAGGTAATTAATCGTACCCATGACGATGTAAGTAAATCCCGGGCTTAACTGGGAAT	600	
Qy	601	TGCGTTTGAACCTACAAAGCTAGATGTGTGCGCAGAGGAGGTGAATTCATGTGTAGCAG	660	

```

Db      |TGCCTTTAACTACAAAGCTAGAGTGTGCGAGAGAGTGGAAATTCATCTGTAGCAG 660
Qy      |TGAATAATGCGTAGAGATATGAAAGAAATCATGATGGCGAGAGCTCTGGGTTAACT 720
Db      |TGAATAATGCGTAGAGATATGAAAGAAATCATGATGGCGAGAGCTCTGGGTTAACT 720
Qy      |GACGCTCATGCAAGAAAGCGTGGAGAGCAAAAGATTAGATACCCTGTAGTCCAGCGC 780
Db      |GACGCTCATGCAAGAAAGCGTGGAGAGCAAAAGATTAGATACCCTGTAGTCCAGCGC 780
Qy      |CTAAACGATGTCAACTAGTTGTTGGGCTTATTAGGCTTGTATGAAAGCTAACCGGTGA 840
Db      |CTAAACGATGTCAACTAGTTGTTGGGCTTATTAGGCTTGTATGAAAGCTAACCGGTGA 840
Qy      |AGTTGACCGCTGGGAGATACGCTCGCAAGATTAAACTCAAGGAATTGACGGGAGACC 900
Db      |AGTTGACCGCTGGGAGATACGCTCGCAAGATTAAACTCAAGGAATTGACGGGAGACC 900
Qy      |GCACAAAGCGGTGGATTTATGTGATTTAATTCGATGCAACCGGAAAAAACCCTTACCTT 960
Db      |GCACAAAGCGGTGGATTTATGTGATTTAATTCGATGCAACCGGAAAAAACCCTTACCTT 960
Qy      |GACATGTAGCGAATTTTCTAGAGATAGATTAGTCTTGGGAAACGCTAACACAGGTGCTG 1020
Db      |GACATGTAGCGAATTTTCTAGAGATAGATTAGTCTTGGGAAACGCTAACACAGGTGCTG 1020
Qy      |CATGGCTGTGCTGACGCTGCTGTGTGAGATGTTGGGTTAAGTCCGCAACGAGCCGCAAC 1080
Db      |CATGGCTGTGCTGACGCTGCTGTGTGAGATGTTGGGTTAAGTCCGCAACGAGCCGCAAC 1080
Qy      |CTTGTCATTTAATTTGTCATCATTTTGGTGGGCACTTTAAGAGACTGCGCGTACAAACCG 1140
Db      |CTTGTCATTTAATTTGTCATCATTTTGGTGGGCACTTTAAGAGACTGCGCGTACAAACCG 1140
Qy      |GAGGAAGGTGGGAGTGAAGTCAAGTCTCTCATGAGCCCTTATGGGTAGGGCTTACACGTAA 1200
Db      |GAGGAAGGTGGGAGTGAAGTCAAGTCTCTCATGAGCCCTTATGGGTAGGGCTTACACGTAA 1200
Qy      |TACAAATGGCGCGTACAGAGAGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAAAGCGCG 1260
Db      |TACAAATGGCGCGTACAGAGAGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAAAGCGCG 1260
Qy      |TCGTAGTCCGATTCGGAATCTGCAACTGCACTCCGTAAGTGGTAATGCTGTATATGCC 1320
Db      |TCGTAGTCCGATTCGGAATCTGCAACTGCACTCCGTAAGTGGTAATGCTGTATATGCC 1320
Qy      |GGATAGAGATGTCGGGAGTATACGTTCCGGGATCTTGTATACACCGCCGCTCACACAT 1380
Db      |GGATAGAGATGTCGGGAGTATACGTTCCGGGATCTTGTATACACCGCCGCTCACACAT 1380
Qy      |GGAGATGGGTTTACACAGAAAGCAGGATGCTAACCGTAAAGAGGGCGCTTGCACGGTGA 1440
Db      |GGAGATGGGTTTACACAGAAAGCAGGATGCTAACCGTAAAGAGGGCGCTTGCACGGTGA 1440
Qy      |GATTATGATCTGGGCTG 1457
Db      |GATTATGATCTGGGCTG 1457

```

```

; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type A R7clone140 16S rDNA
US-10-659-983A-1

Query Match      99.7%; Score 1452.2; DB 7; Length 1457;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      |1 ATTGAAGCTGGCGGATGCTTTACATGCAAGTCGAACGGGACGACGAGATGCTTGAT 60
Db      |1 ATTGAAGCTGGCGGATGCTTTACATGCAAGTCGAACGGGACGACGAGATGCTTGAT 60
Qy      |61 CTGGTGGGAGTGGCGGACGGGTGATATGATGATGCAACGTTATCCAGAAAGGGGGTAA 120
Db      |61 CTGGTGGGAGTGGCGGACGGGTGATATGATGATGCAACGTTATCCAGAAAGGGGGTAA 120
Qy      |121 ACGCATCGAAGATGCTTAATACCGCATATCTCTTAAGAGAGAAAGCAGGGGATCCGAAA 180
Db      |121 ACGCATCGAAGATGCTTAATACCGCATATCTCTTAAGAGAGAAAGCAGGGGATCCGAAA 180
Qy      |121 ACGCATCGAAGATGCTTAATACCGCATATCTCTTAAGAGAGAAAGCAGGGGATCCGAAA 180
Db      |121 ACGCATCGAAGATGCTTAATACCGCATATCTCTTAAGAGAGAAAGCAGGGGATCCGAAA 180
Qy      |181 GACCTTGGCTTTTGGAGCGGCGCATGTCGATTAGTATGTTGGTGGGTTAAAGGCTTAC 240
Db      |181 GACCTTGGCTTTTGGAGCGGCGCATGTCGATTAGTATGTTGGTGGGTTAAAGGCTTAC 240
Qy      |241 CAAGCGACGATCACTAGTATGTTGCTGAGAGAGACACACGACACTGGGACTGAGACAG 300
Db      |241 CAAGCGACGATCACTAGTATGTTGCTGAGAGAGACACACGACACTGGGACTGAGACAG 300
Qy      |301 GCCCAGATCTCTTACCGGAGGCGACGATGGGGAAATTTTGAACAAATGGGGCAAGCCTTATC 360
Db      |301 GCCCAGATCTCTTACCGGAGGCGACGATGGGGAAATTTTGAACAAATGGGGCAAGCCTTATC 360
Qy      |361 CAGCAATGCGCGGTAGAGAGAAAGGCGCTTGGGTTGTAAGCTTTCAAGTCGAGAAAGA 420
Db      |361 CAGCAATGCGCGGTAGAGAGAAAGGCGCTTGGGTTGTAAGCTTTCAAGTCGAGAAAGA 420
Qy      |421 AAAGTTACGGTAAATTAATCGTGAACCATGACCGTATCGACAGAAAGACCGGCTAAC 480
Db      |421 AAAGTTACGGTAAATTAATCGTGAACCATGACCGTATCGACAGAAAGACCGGCTAAC 480
Qy      |481 TACGTCCAGACAGCCCGGTAATACGTAGGGGTGCAAGCGTTAATGGAATTAATCTGGGGT 540
Db      |481 TACGTCCAGACAGCCCGGTAATACGTAGGGGTGCAAGCGTTAATGGAATTAATCTGGGGT 540
Qy      |541 AAAGGTGCGGACGGGCGCTTGTATGATGTAATGTAATCCCGGGCTTAACTTGGGAAT 600
Db      |541 AAAGGTGCGGACGGGCGCTTGTATGATGTAATGTAATCCCGGGCTTAACTTGGGAAT 600
Qy      |601 TGCCTTTAACTACAAAGCTAGAGTGTGCGAGAGAGTGGAAATTCATCTGTAGCAG 660
Db      |601 TGCCTTTAACTACAAAGCTAGAGTGTGCGAGAGAGTGGAAATTCATCTGTAGCAG 660
Qy      |661 TGAATAATGCGTAGAGATATGAAAGAAATCATGATGGCGAGAGCTCTGGGTTAACT 720
Db      |661 TGAATAATGCGTAGAGATATGAAAGAAATCATGATGGCGAGAGCTCTGGGTTAACT 720
Qy      |GACGCTCATGCAAGAAAGCGTGGAGAGCAAAAGATTAGATACCCTGTAGTCCAGCGC 780
Db      |GACGCTCATGCAAGAAAGCGTGGAGAGCAAAAGATTAGATACCCTGTAGTCCAGCGC 780
Qy      |CTAAACGATGTCAACTAGTTGTTGGGCTTATTAGGCTTGTATGAAAGCTAACCGGTGA 840
Db      |CTAAACGATGTCAACTAGTTGTTGGGCTTATTAGGCTTGTATGAAAGCTAACCGGTGA 840

```

```

RESULT 6
US-10-659-983A-1
; Sequence 1, Application US/10659983A
; Publication No. US20040157313A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneq, Timothy A
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284779
; CURRENT APPLICATION NUMBER: US/10/659,983A
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217

```

Db 781 CTAACGATGTCACTAGTTGTTGGGCTTATTAGGCTTGTTACGAACTAACGGCTGA 840  
QY 841 AGTTGACCCGCTGGGGAGTACGGTCCGCAAGATTAACTCAAGAAATTGACGGGAGACC 900  
Db 841 AGTTGACCCGCTGGGGAGTACGGTCCGCAAGATTAACTCAAGAAATTGACGGGAGACC 900  
QY 901 GCACAAAGCGGTGATTAATGATTAATTAATGATGCAACGCGGAAACCTTACTACCTT 960  
Db 901 GCACAAAGCGGTGATTAATGATTAATTAATGATGCAACGCGGAAACCTTACTACCTT 960  
QY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTCTTGGGAAAGCTTACACAGGTGCT 1020  
Db 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTCTTGGGAAAGCTTACACAGGTGCT 1020  
QY 1021 CATGGCTGTGTCAGCTCGGTGTGCTGATGATGTTGGGTTAAAGTCCCGCAAGACGGCAAC 1080  
Db 1021 CATGGCTGTGTCAGCTCGGTGTGCTGATGATGTTGGGTTAAAGTCCCGCAAGACGGCAAC 1080  
QY 1081 CTGTGATTAATTGCCATCAATTTGGTGGGCACTTTAATGACACTGCGGTGACAAACCG 1140  
Db 1081 CTGTGATTAATTGCCATCAATTTGGTGGGCACTTTAATGACACTGCGGTGACAAACCG 1140  
QY 1141 GAGGAAGGTGGGATGACGTCGAAGTCTCATGGCCCTTATGGGTAAGGGCTTCAACGTTA 1200  
Db 1141 GAGGAAGGTGGGATGACGTCGAAGTCTCATGGCCCTTATGGGTAAGGGCTTCAACGTTA 1200  
QY 1201 TACAAATGGCGGCTACAGAGGGTTGCCAACCAGCGAGGGGAGCTAATCTGAGAAAGCGC 1260  
Db 1201 TACAAATGGCGGCTACAGAGGGTTGCCAACCAGCGAGGGGAGCTAATCTGAGAAAGCGC 1260  
QY 1261 TCGTAGTCCGGAATCGAAGTCTGCAACTGACTCCGTAAGTCCGTAATGCTGCT 1320  
Db 1261 TCGTAGTCCGGAATCGAAGTCTGCAACTGACTCCGTAAGTCCGTAATGCTGCT 1320  
QY 1321 GGAATCAGCATGTGCGCGGTGAATACGTTCCGCGGTCTTGTACACACCCCGCTCAACAT 1380  
Db 1321 GGAATCAGCATGTGCGCGGTGAATACGTTCCGCGGTCTTGTACACACCCCGCTCAACAT 1380  
QY 1381 GGGAGTGGGTTTCCAGCAAGCAGGTAGTCTAACCGTAAGGAGGGCGCTTGCACGGTGA 1440  
Db 1381 GGGAGTGGGTTTCCAGCAAGCAGGTAGTCTAACCGTAAGGAGGGCGCTTGCACGGTGA 1440  
QY 1441 GATTCACTGCTGGGCTG 1457  
Db 1441 GATTCACTGCTGGGCTG 1457

RESULT 7  
US-10-659-948A-20  
; Sequence 20, Application US/10659948A  
; Publication No. US20040101946A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovaneq, Timothy A  
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-294309  
; CURRENT APPLICATION NUMBER: US/10/659,948A  
; CURRENT FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573,684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386,217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,219  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 20  
; LENGTH: 1491  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: N. Aestuarii-like AOB B16clone57 16S rDNA

US-10-659-948A-20  
Query Match 95.4%; Score 1390; DB 7; Length 1491;  
Best Local Similarity 97.5%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 35; Indels 1; Gaps 1;  
QY 1 ATTGAACGCTGGCGGATGCTTTTACATGCAAGTCGAACGGCAGACCGATGCTTGCAT 60  
Db 18 ATTGAACGCTGGCGGATGCTTTTACATGCAAGTCGAACGGCAGACCGATGCTTGCAT 77  
QY 61 CTGTGTGCGAGTGGCGGAGCGGCTGAGTAATGCAATGCAACGTAATCCAGAAAGGGGGCTA 120  
Db 78 CTGTGTGCGAGTGGCGGAGCGGCTGAGTAATGCAATGCAACGTAATCCAGAAAGGGGGCTA 137  
QY 121 ACCGATCGAAAGTGTGCTAATACCGCATATCTCTTAAGAGGAAAGCGGGGATGAA 180  
Db 138 ACCGATCGAAAGTGTGCTAATACCGCATATCTCTTAAGAGGAAAGCGGGGATGAA 197  
QY 181 GACCTTGGCGCTTTTGAAGCGGCGGATGCTGATTAAGCTGTTGGGTTAAAGGCTTAC 240  
Db 198 GACCTTGTGCTTTTGAAGCGGCGGATGCTGATTAAGCTGTTGGGTTAAAGGCTTAC 257  
QY 241 CAAGCGAGCATGATGATGTTGCTGAGAGAGCAGCAAGCACTGGGACTGAGACAG 300  
Db 258 CAAGCGAGCATGATGATGTTGCTGAGAGAGCAGCAAGCACTGGGACTGAGACAG 317  
QY 301 GCCGAGCTCTTACGGGAGGAGCAAGTGGGGAAATTTTGAACAATGGGCGGAGCGTATC 360  
Db 318 GCCGAGCTCTTACGGGAGGAGCAAGTGGGGAAATTTTGAACAATGGGCGGAGCGTATC 377  
QY 361 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCCGCTTGAAGCTCTTTCAGTGAAGAGA 420  
Db 378 CAGCAATGCCGCGTGAAGTGAAGAGGCTTCCGCTTGAAGCTCTTTCAGTGAAGAGA 437  
QY 421 AAAGTTACGGTAATTAATGTAAGTACCATGACCGTATGACAGAAAGAGCAGCGCTAAC 480  
Db 438 AAAGTTACGGTAATTAATGTAAGTACCATGACCGTATGACAGAAAGAGCAGCGCTAAC 497  
QY 481 TACGTGCGAGCAGCGCGGTGAATACGTAAGGTGCAAGCGTTAATCGGAATTAATCGGGCGT 540  
Db 498 TACGTGCGAGCAGCGCGGTGAATACGTAAGGTGCAAGCGTTAATCGGAATTAATCGGGCGT 557  
QY 541 AAAGGTGCGCAGCGCGGCTTGAAGTCAATGTAAGTCAATCCCGGCTTAACTGGGAAT 600  
Db 558 AAAGGTGCGCAGCGCGGCTTGAAGTCAATGTAAGTCAATCCCGGCTTAACTGGGAAT 617  
QY 601 TCGCTTGAAGACTCAAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTACAG 660  
Db 618 TCGCTTGAAGACTCAAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTACAG 677  
QY 661 TGAATGCGTAGAGATATGAAAGACATGATGCGAAGGCAAGCTCTCGGCTTAACTACT 720  
Db 678 TGAATGCGTAGAGATATGAAAGACATGATGCGAAGGCAAGCTCTCGGCTTAACTACT 737  
QY 721 GACGCTATGACCAAAAGCGTGGGAGCAAAACAGATTAAGTATACCTGGTATGCAAGCC 780  
Db 738 GACGCTATGACCAAAAGCGTGGGAGCAAAACAGATTAAGTATACCTGGTATGCAAGCC 797  
QY 781 CTAACGATGCACTAGTTGTTGGGCTTATTAAGCTTGTGTAACGAAGCTAACGCGTGA 840  
Db 798 CTAACGATGCACTAGTTGTTGGGCTTATTAAGCTTGTGTAACGAAGCTAACGCGTGA 857  
QY 841 AGTTGACCCGCTGGGGAGTACGGTCCGCAAGTTAAACTCAAGGAATTGACGGGAGACC 900  
Db 858 AGTTGACCCGCTGGGGAGTACGGTCCGCAAGTTAAACTCAAGGAATTGACGGGAGACC 917  
QY 901 GCACAAAGCGGTGATTAATGATTAATTAATGATGCAACGCGGAAACCTTACTACCTT 960  
Db 918 GCACAAAGCGGTGATTAATGATTAATTAATGATGCAACGCGGAAACCTTACTACCTT 977  
QY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTCTTGGGAAAGCTTACACAGGTGCT 1019  
Db 978 GACATGTAGCGAATTTTCTAGAGATAGATTAGTCTTGGGAAAGCTTACACAGGTGCT 1037

QY 1020 GCATGCTGTCTGACGCTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCCGCAAC 1079  
DB 1038 GCATGCTGTCTGACGCTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCCGCAAC 1097  
QY 1080 CCTTGCAATTAATGTCATCATTTGGTGGGCACTTTAATGAGACTGCCGGTGAACAAC 1139  
DB 1098 CCTTGCAATTAATGTCATCATTTGGTGGGCACTTTAATGAGACTGCCGGTGAACAAC 1157  
QY 1140 GGAGGAGGTGGGATGAGCTCAAGTCTCATGAGCCCTTAATGGGTAAGGCTTCAACGTA 1199  
DB 1158 GGAGGAGGTGGGATGAGCTCAAGTCTCATGAGCCCTTAATGGGTAAGGCTTCAACGTA 1217  
QY 1200 ATTCATATGCGCGCTACAGAGGGTGGCCAAACCCCGAGGGGAGCTAATCTCAGAAAGCCG 1259  
DB 1218 ATTCATATGCGCGCTACAGAGGGTGGCCAAACCCCGAGGGGAGCTAATCTCAGAAAGCCG 1277  
QY 1260 GTCTGATGCTCGGATCGGAGTCTGCAACTGCACTCCGTAAGTGGAGTGGTGAATGCG 1319  
DB 1278 GTCTGATGCTCGGATCGGAGTCTGCAACTGCACTCCGTAAGTGGAGTGGTGAATGCG 1337  
QY 1320 CGGATCAGATGTCTGCGGTGAATACGTTCCCGGCTCTTGACACACCGCCGTCACACCA 1379  
DB 1338 CGGATCAGATGTCTGCGGTGAATACGTTCCCGGCTCTTGACACACCGCCGTCACACCA 1397  
QY 1380 TGGGATGGGTTTACACCAAGCAAGTATCTTAACCTAAGAGGGGCGCTTGACAGGTG 1439  
DB 1398 TGGGATGGGTTTACACCAAGCAAGTATCTTAACCTAAGAGGGGCGCTTGACAGGTG 1457  
QY 1440 AGATTGATGCTGGGGTG 1457  
DB 1458 AGATTGATGCTGGGGTG 1475

RESULT 8  
US-10-659-980a-20

/ Sequence 20, Application US/10659980A  
/ Publication No. US2004010613A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Hovaneq, Timothy A  
/ TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria  
/ FILE REFERENCE: 81289-284781  
/ CURRENT APPLICATION NUMBER: US/10/659, 980A  
/ CURRENT FILING DATE: 2003-09-10  
/ PRIOR APPLICATION NUMBER: US 09/573,684  
/ PRIOR FILING DATE: 2000-05-19  
/ PRIOR APPLICATION NUMBER: US 60/386,217  
/ PRIOR FILING DATE: 2002-09-19  
/ PRIOR APPLICATION NUMBER: US 60/386,218  
/ PRIOR FILING DATE: 2002-09-19  
/ PRIOR APPLICATION NUMBER: US 60/386,219  
/ PRIOR FILING DATE: 2002-09-19  
/ NUMBER OF SEQ ID NOS: 23  
/ SOFTWARE: PatentIn version 3.2  
/ SEQ ID NO 20  
/ LENGTH: 1491  
/ TYPE: DNA  
/ ORGANISM: Unknown  
/ FEATURE:  
/ OTHER INFORMATION: N. Aestuarii-like AOB BFl6c1one57 16S rDNA  
US-10-659-980a-20

Query Match 95.4%; Score 1390; DB 7; Length 1491;  
Best Local Similarity 97.5%; Pred. No. 0;  
Matches 1422; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 1 ATTTGAACGCTGGGCGATGTTTACATGCAAGTGAACGCGACGCGATGCTTGAT 60  
DB 18 ATTTGAACGCTGGGCGATGTTTACATGCAAGTGAACGCGACGCGATGCTTGAT 77  
QY 61 CTGGTGGCGAGTGGCGGACGGGTGAGTATGCAATCGGAAAGTATCCAGAAAGAGGGGGTA 120  
DB 78 CTGGTGGCGAGTGGCGGACGGGTGAGTATGCAATCGGAAAGTATCCAGAAAGTGGGGATA 137

QY 121 AGCGATCGAAAGATGTCGTAATACCGCATATATCTTAAGSAGAAAGCAGGGGATCGAA 180  
DB 138 AGCGATCGAAAGATGTCGTAATACCGCATATATCTTAAGSAGAAAGCAGGGGATCGAA 197  
QY 181 GACCTTGCGCTTTTGGAGCGGCGATGTCGTAATGATGTTGGTGGGTTAAAGGCTTAC 240  
DB 198 GACCTTGCGCTTTTGGAGCGGCGATGTCGTAATGATGTTGGTGGGTTAAAGGCTTAC 257  
QY 241 CAAGCGACGATCATGATGTTGGTCTGAGAGGACGACACGACCATCTGGGACTGACACG 300  
DB 258 CAAGCGACGATCATGATGTTGGTCTGAGAGGACGACACGACCATCTGGGACTGACACG 317  
QY 301 GCCGAGCTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACATAGGCGCAAGCTGATC 360  
DB 318 GCCGAGCTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACATAGGCGCAAGCTGATC 377  
QY 361 CAGCAATGCGCGGTGATGTAAGAGGACCTTCCGGTTGTAAAGCTCTTTCAGTCGAGAGA 420  
DB 378 CAGCAATGCGCGGTGATGTAAGAGGACCTTCCGGTTGTAAAGCTCTTTCAGTCGAGAGA 437  
QY 421 AAAGTTACGTTAATATCTGACCCATGACGCTATGACAGAAAGAACACCGGTTAC 480  
DB 438 AAAGTTACGTTAATATCTGACCCATGACGCTATGACAGAAAGAACACCGGTTAC 497  
QY 481 TAAGTCCAGACGCGCGGTTAATAGTGGGTGCAAGCGTTAATCGAATTACTGGGCGT 540  
DB 498 TAAGTCCAGACGCGCGGTTAATAGTGGGTGCAAGCGTTAATCGAATTACTGGGCGT 557  
QY 541 AAAGGTGCGCAGCGCGCTTGTAAATGATGTAATCCCGGCGTTAACTTGGGAAT 600  
DB 558 AAAGGTGCGCAGCGCGCTTGTAAATGATGTAATCCCGGCGTTAACTTGGGAAT 617  
QY 601 TCGCTTTGAACTTACAAAGCTTGAATGTCGAGAGGAGTGGAAATTCATGTCGACG 660  
DB 618 TCGCTTTGAACTTACAAAGCTTGAATGTCGAGAGGAGTGGAAATTCATGTCGACG 677  
QY 661 TGAATGCGTGAATATGTAAGAAACATGATGCGGAGGAGCGAGCTCTGGGTTAACAT 720  
DB 678 TGAATGCGTGAATATGTAAGAAACATGATGCGGAGGAGCGAGCTCTGGGTTAACAT 737  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGATACCTGGTACTGACGCGC 780  
DB 738 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGATACCTGGTACTGACGCGC 797  
QY 781 CTAAACGATGTCATGATGTTGGTGGGCTTATTAAGCTTGGTATGCAAGCTAACCGGTGA 840  
DB 798 CTAAACGATGTCATGATGTTGGTGGGCTTATTAAGCTTGGTATGCAAGCTAACCGGTGA 857  
QY 841 AGTTGACGCGCTGGGGAATACGCTGCAAGATTAATACTCAAGAAATGACGCGGAGCC 900  
DB 858 AGTTGACGCGCTGGGGAATACGCTGCAAGATTAATACTCAAGAAATGACGCGGAGCC 917  
QY 901 GCAACAAGCGGTGATTAATGATGATTAATTCATGCAACGCGAAACCTTACTACCTT 960  
DB 918 GCAACAAGCGGTGATTAATGATGATTAATTCATGCAACGCGAAACCTTACTACCTT 977  
QY 961 GACATGACGAAATTTTCTAGAGATGATTAAGTG-CTTGGGAAACCTTAAACAAGTGTCT 1019  
DB 978 GACATGACGAAATTTTCTAGAGATGATTAAGTGAGTGGGAAACCTTAAACAAGTGTCT 1037  
QY 1020 GCATGCTGTCTGACGCTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCCGCAAC 1079  
DB 1038 GCATGCTGTCTGACGCTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCCGCAAC 1097  
QY 1080 CCTTGCAATTAATGTCATCATTTGGTGGGCACTTTAATGAGACTGCCGGTGAACAAC 1139  
DB 1098 CCTTGCAATTAATGTCATCATTTGGTGGGCACTTTAATGAGACTGCCGGTGAACAAC 1157  
QY 1140 GGAGGAGGTGGGATGAGCTCAAGTCTCATGAGCCCTTAATGGGTAGGCTTCAACGTA 1199  
DB 1158 GGAGGAGGTGGGATGAGCTCAAGTCTCATGAGCCCTTAATGGGTAGGCTTCAACGTA 1217

QY 1200 ATACAAATGCGCGCTACAGAGGGTTGCTCAACCCGCGAGGGGAGCTTAATCTCAGAAAGCCG 1259  
|  
|  
|  
Db 1218 ATACAAATGCGCGCTACAGAGGGTTGCTCAACCCGCGAGGGGAGCTTAATCTCAGAAAGCCG 1277  
|  
|  
|  
QY 1260 GTCTGATGCTCCGATTCGGAATCTGCACTCGGATGAGGATGCGGATGCTGATATCG 1319  
|  
|  
|  
Db 1278 GTCTGATGCTCCGATTCGGAATCTGCACTCGGATGAGGATGCGGATGCTGATATCG 1337  
|  
|  
|  
QY 1320 CGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCCGCCGTCAACCA 1379  
|  
|  
|  
Db 1338 CGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCCGCCGTCAACCA 1397  
|  
|  
|  
QY 1380 TGGAGTGGGTTTCAACGAAAGAGGTATCTTAACCGTAAAGAGGCGCTTGCCACGCTG 1439  
|  
|  
|  
Db 1398 TGGAGTGGGTTTCAACGAAAGAGGTATCTTAACCGTAAAGAGGCGCTTGCCACGCTG 1457  
|  
|  
|  
QY 1440 AGATTCAATGCTGGGGTG 1457  
|  
|  
|  
Db 1458 AGATTCAATGCTGGGGTG 1475  
|  
|  
|

## RESULT 9

US-10-659-983a-20  
; Sequence 20, Application US/10659983A  
; Publication No. US2004015731A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovaneq, Timothy A  
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-284779  
; CURRENT APPLICATION NUMBER: US/10/659, 983A  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573,684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386,217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,219  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 20  
; LENGTH: 1491  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: N. aestuarii-like AOB B16clones7 16S rDNA  
; US-10-659-983a-20

Query Match 95.4%; Score 1390; DB 7; Length 1491;

Best Local Similarity 97.5%; Pred. No. 0;

Matches 1422; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 1 ATGGAACGCTGGCGGATGCTTTACACATGCAAGTCGAACGGCAGCAACGATGCTTGAT 60  
|  
|  
|  
Db 18 ATGGAACGCTGGCGGATGCTTTACACATGCAAGTCGAACGGCAGCAACGATGCTTGAT 77  
|  
|  
|  
QY 61 CTGTGTGCGATGCGGACGCGGTGATTAATGCAATCGGAACGTATCCGAAGAGGGGGTA 120  
|  
|  
|  
Db 78 CTGTGTGCGATGCGGACGCGGTGATTAATGCAATCGGAACGTATCCGAAGAGGGGGTA 137  
|  
|  
|  
QY 121 ACCGATCGAAGATGTGTCTAATACCGCATATACTCTAAGAGAGAAAGCAGGGAGTCGAAA 180  
|  
|  
|  
Db 138 ACCGATCGAAGATGTGTCTAATACCGCATATACTCTAAGAGAGAAAGCAGGGAGTCGAAA 197  
|  
|  
|  
QY 181 GACCTTGGCTTTTGAAGCGGCGGATGCTGATTAAGCTAAGTTGGTGGGTTAAAGGCTTAC 240  
|  
|  
|  
Db 198 GACCTTGGCTTTTGAAGCGGCGGATGCTGATTAAGCTAAGTTGGTGGGTTAAAGGCTTAC 257  
|  
|  
|  
QY 241 CAAAGCCAGCATCATGATGTTGTCTGAGAGAGACGACGACCACTGGGAGCTGAGACAG 300  
|  
|  
|  
Db 258 CAAAGCCAGCATCATGATGTTGTCTGAGAGAGACGACGACCACTGGGAGCTGAGACAG 317  
|  
|  
|

QY 301 GCCCAAGCTCTTAACGGGAGGACAGCATGGGGAAATTTTGGACATATGGGCGCAAGCTGATC 360  
|  
|  
|  
Db 318 GCCCAAGCTCTTAACGGGAGGACAGCATGGGGAAATTTTGGACATATGGGCGCAAGCTGATC 377  
|  
|  
|  
QY 361 CAGCAATGCGCGGTGAGTGAAGAGAGCCCTTGGGTTGTAAGCTCTTTCACTGAGAGAGA 420  
|  
|  
|  
Db 378 CAGCAATGCGCGGTGAGTGAAGAGAGCCCTTGGGTTGTAAGCTCTTTCACTGAGAGAGA 437  
|  
|  
|  
QY 421 AAAGTTACGGTAATTAATCTGACCCCATGACGGTATCCGACAGAAAGCAGCGGGCTAAC 480  
|  
|  
|  
Db 438 AAAGTTACGGTAATTAATTAATCAATTAATGACGGTATCCGACAGAAAGCAGCGGGCTAAC 497  
|  
|  
|  
QY 481 TACGTCCAGACGCGCGGTAAATAGTAGGGGTCAAGCGTTAATCGGAATTACTGGGCGT 540  
|  
|  
|  
Db 498 TACGTCCAGACGCGCGGTAAATAGTAGGGGTCAAGCGTTAATCGGAATTACTGGGCGT 557  
|  
|  
|  
QY 541 AAAGGTGGCGCAGGCGCGCTTGTAACTGATGATGTAATCCCGGGCTTAACTGGGAAAT 600  
|  
|  
|  
Db 558 AAAGGTGGCGCAGGCGCGCTTGTAACTGATGATGTAATCCCGGGCTTAACTGGGAAAT 617  
|  
|  
|  
QY 601 TCGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGGAGGTGAATTCATGTTAGCAG 660  
|  
|  
|  
Db 618 TCGGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGGAGGTGAATTCATGTTAGCAG 677  
|  
|  
|  
QY 661 TGAATTCGCTAGATATGGAAGACATGATGCGGACAGCTCTCTGGTTAACACT 720  
|  
|  
|  
Db 678 TGAATTCGCTAGATATGGAAGACATGATGCGGACAGCTCTCTGGTTAACACT 737  
|  
|  
|  
QY 721 GACGCTCATGACGAAAGCGTGGGAGACAAACAGATTAGATACCTGATGTCAGGCC 780  
|  
|  
|  
Db 738 GACGCTCATGACGAAAGCGTGGGAGACAAACAGATTAGATACCTGATGTCAGGCC 797  
|  
|  
|  
QY 781 CTAAGCATGTCACTAGTTGTTGGGCTTTATTAAGCTTGTAACGAACTTAACGCTGA 840  
|  
|  
|  
Db 798 CTAAGCATGTCACTAGTTGTTGGGCTTTATTAAGCTTGTAACGAACTTAACGCTGA 857  
|  
|  
|  
QY 841 AGTTGACCGGCTGGGAGATACGCTCCCAAGATTAAACCTAAAGAAATGACGGGAGCC 900  
|  
|  
|  
Db 858 AGTTGACCGGCTGGGAGATACGCTCCCAAGATTAAACCTAAAGAAATGACGGGAGCC 917  
|  
|  
|  
QY 901 GCAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAATACTTACCTACCTT 960  
|  
|  
|  
Db 918 GCAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAATACTTACCTACCTT 977  
|  
|  
|  
QY 961 GACATGATGCAATTTTCTAGAGATGATTAATGTCGCGGAAACGCTTAACAGAGTCT 1019  
|  
|  
|  
Db 978 GACATGATGCAATTTTCTAGAGATGATTAATGTCGCGGAAACGCTTAACAGAGTCT 1037  
|  
|  
|  
QY 1020 GCATGCTGTCTGATGCTGCTGCTGCTGATGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1079  
|  
|  
|  
Db 1038 GCATGCTGTCTGATGCTGCTGCTGATGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1097  
|  
|  
|  
QY 1080 CTTGTGATTAATTTGCAATCATTTGTTGGGCACTTTAATAGACCTGCGGATGCAAAAC 1139  
|  
|  
|  
Db 1098 CTTGTGATTAATTTGCAATCATTTAATTTAGTGGGCACTTTAATAGACCTGCGGATGCAAAAC 1157  
|  
|  
|  
QY 1140 GAGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199  
|  
|  
|  
Db 1158 GAGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1217  
|  
|  
|  
QY 1200 ATACAAATGCGCGCTGACAGAGGTTGCGAACCCCGGAGAGGGAGCTTAATCTCAGAAAGCGC 1259  
|  
|  
|  
Db 1218 ATACAAATGCGCGCTGACAGAGGTTGCGAACCCCGGAGAGGGAGCTTAATCTCAGAAAGCGC 1277  
|  
|  
|  
QY 1260 GTCTGATGCTCCGATTCGGAATCTGCACTCGGATGAGGATGCGGATGCTGATATCG 1319  
|  
|  
|  
Db 1278 GTCTGATGCTCCGATTCGGAATCTGCACTCGGATGAGGATGCGGATGCTGATATCG 1337  
|  
|  
|  
QY 1320 CGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCCGCCGTCAACCA 1379  
|  
|  
|  
Db 1338 CGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCCGCCGTCAACCA 1397  
|  
|  
|  
QY 1380 TGGAGTGGGTTTCAACGAAAGAGGTATCTTAACCGTAAAGAGGCGCTTGCCACGCTG 1439  
|  
|  
|

```
Db      1398 TGGAGTGGGTTTACACGAGACGATAGTCTAACCGTAAGAGGGCGTTTGCCACGGTG 1457
QY      1440 AGATTGATGACTGGGGTG 1457
Db      1458 AGATTGATGACTGGGGTG 1475

RESULT 10
US-10-659-948A-19
; Sequence 19, Application US/10659948A
; Publication No. US20040101946A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneec, Timothy A
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-294309
; CURRENT APPLICATION NUMBER: US/10/659,948A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR FILING DATE: 2002-09-19
; SOFTWARE: PatentIn version 3.2
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 19
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: N. Aestuarii-like AOB Paclone31 16S rDNA
US-10-659-948A-19

Query Match      93.2%; Score 1357.6; DB 7; Length 1494;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1401; Conservative 0; Mismatches 54; Indels 1; Gaps 1;
```

```
Db      499 TACGTGCCAGACGCCGCGTAAATAGTAGGGTGCAGAGGTTAATCGAATTACTGGCGGT 558
QY      541 AAAAGGTGCGACGGGCGCTTTGAAGTCAGATGTGAATCCCCGGGCTTAACCTGGGAAT 600
Db      559 AAAAGGTGCGACGGGCGCTTTGAAGTCAGATGTGAATCCCCGGGCTTAACCTGGGAAT 618
QY      601 TGGCGTTTGAACCTAACAAAGCTAGAGTGGCGAGAGGAGTGAATTCATGTGTAGCAG 660
Db      619 TCGGTTTGAACCTAACAAAGCTAGAGTGGCGAGAGGAGTGAATTCATGTGTAGCAG 678
QY      661 TGAATGCGTAGAGATTAAGAAACAATCGATGCGAAGGACGCTCTCGGTTAACT 720
Db      679 TGAATGCGTAGAGATTAAGAAACAATCGATGCGAAGGACGCTCTCGGTTAACT 738
QY      721 GACGCTCATGACGAAAGCGTGGGAGCAACAAGATTAGATACCTGTAGTCCAGCC 780
Db      739 GACGCTCATGACGAAAGCGTGGGAGCAACAAGATTAGATACCTGTAGTCCAGCC 798
QY      781 CTAAACGATGTCAACTAGTTGTTGGCCCTTAATAGGCTTGTGAAGCTAACCGGTGA 840
Db      799 CTAAACGATGTCAACTAGTTGTTGGCCCTTAATAGGCTTGTGAAGCTAACCGGTGA 858
QY      841 AGTTGACCGCTGGGAGTAGACGTCGCAAGATTAAACTCAAGGAATTGACGGGAGCC 900
Db      859 AGTTGACCGCTGGGAGTAGACGTCGCAAGATTAAACTCAAGGAATTGACGGGAGCC 918
QY      901 GCACAAGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAAACTTACCTACCTT 960
Db      919 GCACAAGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAAACTTACCTACCTT 978
QY      961 GACATGTAGCGCAATTTTCTAGAGATAGATTAGT-CTTCGGGAGAGCTTAACAGGTGCT 1019
Db      979 GACATGTAGCGCAATTTTCTAGAGATAGATTAGT-CTTCGGGAGAGCTTAACAGGTGCT 1038
QY      1020 GCATGGCTGTCGTCACTGCTGCTGCTGAGATGTTGGTTAACTCCGCAACGAGCGCAAC 1079
Db      1039 GCATGGCTGTCGTCACTGCTGCTGAGATGTTGGTTAACTCCGCAACGAGCGCAAC 1098
QY      1080 CCTTGTCAATTAATGCCATCAATTTGGTGGGCACTTTAATGAGACTCCCGGTGACAAACC 1139
Db      1099 CCTTGTCAATTAATGCCATCAATTTGGTGGGCACTTTAATGAGACTCCCGGTGACAAACC 1158
QY      1140 GAGGAGAGGTGGGAGTAGAGTGAATCTTCATAGGCTTATGAGGCTTACACGTA 1199
Db      1159 GAGGAGAGGTGGGAGTAGAGTGAATCTTCATAGGCTTATGAGGCTTACACGTA 1218
QY      1200 ATACAATGCGCGGTACAGAGGGTTGCCAACCCGAGGGGAGACTTAATCTCAGAAAGCGC 1259
Db      1219 ATACAATGCGCGGTACAGAGGGTTGCCAACCCGAGGGGAGACTTAATCTCAGAAAGCGC 1278
QY      1260 GTCTGATGTCGGATCGGAGTCTGCAACTGCACTCCGTGAAGTGGGAATCTGATATCG 1319
Db      1279 GTCTGATGTCGGATCGGAGTCTGCAACTGCACTCCGTGAAGTGGGAATCTGATATCG 1338
QY      1320 CGGATCAGCATGTCCGCGTGAATAGTTTCCGGGCTTGTATACACCGCGCTCACCA 1379
Db      1339 CGGATCAGCATGTCCGCGTGAATAGTTTCCGGGCTTGTATACACCGCGCTCACCA 1398
QY      1380 TGGAGTGGGTTTCAACAGAACAGGTAGTCTTAACGTTAAGGAGGAGCTTCCACGGTG 1439
Db      1399 TGGAGTGGGTTTCAACAGAACAGGTAGTCTTAACGTTAAGGAGGAGCTTCCACGGTG 1458
QY      1440 AGATTGATGACTGGGG 1455
Db      1459 TGGTCAATGACTTGGG 1474

RESULT 11
US-10-659-980A-19
; Sequence 19, Application US/10659980A
; Publication No. US20040106133A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Hovanec, Timothy A
; TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284781
; CURRENT APPLICATION NUMBER: US/10/659,980A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: N. Aestuarii-like AOB Paclone31 16S rDNA
US-10-659-980A-19

Query Match      93.2%; Score 1357.6; DB 7; Length 1494;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1401; Conservative 0; Mismatches 54; Indels 1; Gaps 1;

Qy      1  ATTGAAGCTGGCGGCGGCTTTACATGCAAGTCGAAAGCGGACGACGAGTGGCTTGCAT 60
Db      19  ATTGAAGCTGGCGGCGGCTTTACATGCAAGTCGAAAGCGGACGAGGAGCTTGGCAC 78
Qy      61  CTGGTGGCGAGTGGCGGACGCGGTGATGATGATCGAAGCGTATCGAAGAGGGGGGTA 120
Db      79  CTGGTGGCGAGTGGCGGACGCGGTGATGATGATCGAAGCGTATCGGAGATGGGGGATA 138
Qy      121  ACGCATGAAAGATGTGCTAATACCGCATATCTCTTAAGAGAAAGAGAGGGATTCGAAA 180
Db      139  ACGCATGAAAGATGTGCTAATACCGCATATCTCTTAAGAGAAAGAGAGGGATTCGAAA 198
Qy      181  GACCTTGGCTTTGGAGCGGCGGATGCTGATTAGCTAGTGGGGGTAAAGGCTTAC 240
Db      199  GACCTTGGCTTTGGAGCGGCGGATGCTGATTAGCTAGTGGGGGTAAAGGCTTAC 258
Qy      241  CAAAGCGCGATCAGTAGTGTGCTGAGAGGACGACCACTGGGACTGAGACG 300
Db      259  CAAAGCGCGATCAGTAGTGTGCTGAGAGGACGACCACTGGGACTGAGACG 318
Qy      301  GCCCAGACTCTTACGGGAGGACGAGTGGGAAATTTTGGACAATGGGCGCAAGCTGATC 360
Db      319  GCCCAGACTCTTACGGGAGGACGAGTGGGAAATTTTGGACAACGGGCGCAAGCTGATC 378
Qy      361  CAGCAATCCCGCTGAGTGAAGAAAGGCTTGGGGTGAAGCTTTCAATCGAAGAA 420
Db      379  CAGCAATCCCGCTGAGTGAAGAAAGGCTTGGGGTGAAGCTTTCAATCGAAGAA 438
Qy      421  AAAGGTTACGGTAAATATCTGACCCATGACGATGACAGAGAAAGACCGGCTTAC 480
Db      439  AAAGGTTACGTAATATATCAACTTATGACGGATACCGACAGAAAGACCGGCTTAC 498
Qy      481  TACGTGCAGGACCGCGGTAAATACGTAAGGTGCAAGCGTTAATCGAATTAATCTGGGCT 540
Db      499  TACGTGCAGGACCGCGGTAAATACGTAAGGTGCAAGCGTTAATCGAATTAATCTGGGCT 558
Qy      541  AAAAGGTGCGCAGGCGGCTTGTAGTGAATGTGAATCCCGGGCTTAATCTGGGAAAT 600
Db      559  AAAAGGTGCGCAGGCGGCTTGTAGTGAATGTGAATCCCGGGCTTAATCTGGGAAAT 618
Qy      601  TCGCTTTGAACCTAACAAGCTAGTGTGACAGAGGAGTGAATTCATGTGTAGCAG 660
Db      619  TCGCTTTGAACCTAACAAGCTAGTGTGACAGAGGAGGAGTGAATTCATGTGTAGCAG 678
Qy      661  TGAATGCGTAGATATGAAAGAACTGATGCGGAAAGCGCTCTCTGGGTTAACT 720

```

```

Db      679  TGAATGCGTAGATATGAAAGAACTGATGCGGAAAGCGAGCCCTCTGGGTTAACT 738
Qy      721  GACGCTCATGACGAAAGCTGGGAGGAAACAGATTGATACCTGTGATGTCACGCC 780
Db      739  GACGCTCATGACGAAAGCTGGGAGGAAACAGATTGATACCTGTGATGTCACGCC 798
Qy      781  CTAAAGATGTCAACTAGTGTGGGCTTATTAAGCTTGTGATGACGAACTAAAGCGGTGA 840
Db      799  CTAAAGATGTCAACTAGTGTGGGCTTATTAAGCTTGTGATGACGAACTAAAGCGGTGA 858
Qy      841  AGTTGACCGCTGGGAGTACGGTCCGAAGTTTAACTCAAGAAATGACGGGAGACC 900
Db      859  AGTTGACCGCTGGGAGTACGGTCCGAAGTTTAACTCAAGAAATGACGGGAGACC 918
Qy      901  GCAAGAGCGGTGATTAATGATTAATTCGATGCAACCGGAAACCTTAACCTTACCTT 960
Db      919  GCACAGCGGTGATTAATGATTAATTCGATGCAACCGGAAACCTTAACCTTACCTT 978
Qy      961  GACATGTAGCGAATTTTCTAAGATAGATTAGTG-CTTGGGAAAGCTTAAACAGGTCCT 1019
Db      979  GACATGTAGCGAATTTTCTAAGATAGATTAGTGCTTGGGAAAGCTTAAACAGGTCCT 1038
Qy      1020  GCATGCGTGTGCTCAGCTCGTGTGAGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1079
Db      1039  GCATGCGTGTGCTCAGCTCGTGTGAGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1098
Qy      1080  CTTGTTCATTAATTCGCATCAATTTGGTGGGCACTTAAATGAACTGCGGTGACAAAC 1139
Db      1099  CTTGTTCATTAATTCGCATCAATTTGGTGGGCACTTAAATGAACTGCGGTGACAAAC 1158
Qy      1140  GAGAGAAAGTGGGAGTGAAGTCAAGTCTCATAGGCCCTTAATGAGAGGCTTACACGTA 1199
Db      1159  GAGAGAAAGTGGGAGTGAAGTCAAGTCTCATAGGCCCTTAATGAGAGGCTTACACGTA 1218
Qy      1200  ATACAATGCGCGCTTACAGAGGTTGGCAACCCCGAGGGGAGCTAATCTCAGAAAGCGC 1259
Db      1219  ATACAATGCGCGCTTACAGAGGTTGGCAACCCCGAGGGGAGCTAATCTCAGAAAGCGC 1278
Qy      1260  GTCGTAGTCCGATTCGAGTCTGCAACTGCACTCCGTAAGTGGAAATCGCTAGTATCG 1319
Db      1279  GTCGTAGTCCGATTCGAGTCTGCAACTGCACTCCGTAAGTGGAAATCGCTAGTATCG 1338
Qy      1320  CGATTCAGCATGTCGCGGTGAATACGTTCCGGGCTCTTGACACCGCCGTCACACCA 1379
Db      1339  CGATTCAGCATGTCGCGGTGAATACGTTCCGGGCTCTTGACACCGCCGTCACACCA 1398
Qy      1380  TGGAGTGGGTTTACACGAGAGGATGTAACCTGAAGAGAGGCGCTTGCACGCTG 1439
Db      1399  TGGAGTGGGTTTACACGAGAGGATGTAACCTGAAGAGAGGCGCTTGCACGCTG 1458
Qy      1440  AGATTTCATGACTGGGG 1455
Db      1459  TGCTCATGACTTGGG 1474

RESULT 12
US-10-659-983A-19
; Sequence 19, Application US/10659983A
; Publication No. US20040157313A1
; GENERAL INFORMATION:
; APPLICANT: Hovanec, Timothy A
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284779
; CURRENT APPLICATION NUMBER: US/10/659,983A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19

```



```
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: N. Aestuarii-like AOB P4clone31 16S rDNA
US-10-659-983a-19

Query Match      93.2%; Score 1357.6; DB 7; Length 1494;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1401; Conservative 0; Mismatches 54; Indels 1; Gaps 1;

QY      1 ATTGAAGCTGGCGGCAATGCTTTACATGCAATGCGAAGCGCGACGCGATGCTTGAT 60
      19 ATTGAAGCTGGCGGCAATGCTTTACATGCAATGCGAAGCGCGACGCGATGCTTGAT 78
      61 CTGGTGGCGAGTGGCGGAGCGGAGTATGATCGGAAAGTATCCAGAAAGAGGGGGTGA 120
      79 CTGGTGGCGAGTGGCGGAGCGGAGTATGATCGGAAAGTATCCAGAAAGAGGGGGTGA 138
      121 ACGCATGAAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAA 180
      139 ACGCATGAAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAA 198
      181 GACCTTGGCGTTTGGAGCGGCGCATGCTGATTAAGTAAAGTGGGCTTAAAGGCTTAC 240
      199 GACCTTGGCGTTTGGAGCGGCGCATGCTGATTAAGTAAAGTGGGCTTAAAGGCTTAC 258
      241 CAAGGCGAGCATGATGATGTTGCTGAGAGAGCAGCAGCACACTGGGATCTGAGACAG 300
      259 CAAGGCGAGCATGATGATGTTGCTGAGAGAGCAGCAGCACACTGGGATCTGAGACAG 318
      301 GCCCAGACTCTTACGGAGAGCAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCTGATC 360
      319 GCCCAGACTCTTACGGAGAGCAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCTGATC 378
      361 CAGCAATGCGCGTGAATGAAGAGCGCTTGGGTTGTAAGCTCTTTCAGTGAAGA 420
      379 CAGCAATGCGCGTGAATGAAGAGCGCTTGGGTTGTAAGCTCTTTCAGTGAAGA 438
      421 AAAGGTAAAGGTAAATATGCTGACCATAGCGGATCGAAGAGAAAGACCGGGCTAAC 480
      439 AAAGGTAAAGGTAAATATGCTGACCATAGCGGATCGAAGAGAAAGACCGGGCTAAC 498
      481 TAAGTGCACAGCGCGGTAATATGTAAGGTGCAAGCGTTAATCGAATTAATCTGGGCGT 540
      499 TAAGTGCACAGCGCGGTAATATGTAAGGTGCAAGCGTTAATCGAATTAATCTGGGCGT 558
      541 AAAGGTGCGCAGCGCGGCTTTGTAAGTCAATGTGAATCCCGGGCTTAACTGGGAAT 600
      559 AAAGGTGCGCAGCGCGGCTTTGTAAGTCAATGTGAATCCCGGGCTTAACTGGGAAT 618
      601 TGCGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGAGGTAATCCATGTGTACAG 660
      619 TGCGTTGAAACTACAAAGCTAGAGTGTGGCAGAGGAGGAGGTAATCCATGTGTACAG 678
      661 TGAATATGCTAGAGATATGGAAGAACATCGATGGCGAAGCAGCTCTGGGTTAACTACT 720
      679 TGAATATGCTAGAGATATGGAAGAACATCGATGGCGAAGCAGCTCTGGGTTAACTACT 738
      721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAATACCTCTGGTATCGACGCC 780
      739 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAATACCTCTGGTATCGACGCC 798
      781 CTAAACGATGCAACTATGTTGTTGGGCTTATTAAGTTTGTGTAACGAAGCTAACGCGTGA 840
      799 CTAAACGATGCAACTATGTTGTTGGGCTTATTAAGTTTGTGTAACGAAGCTAACGCGTGA 858
      841 AGTTGACCGGCTGGGAGATACGGTCCGCAATTTAAACTCAAGGAATTTGACGGGACCC 900
      859 AGTTGACCGGCTGGGAGATACGGTCCGCAATTTAACTCAAGGAATTTGACGGGACCC 918
```

```
QY      901 GCACAAAGCGGTGGAATTAATGATTAATTCGATGCAACGCGAANAACCTTACCTACCCCTT 960
      919 GCACAAAGCGGTGGAATTAATGATTAATTCGATGCAACGCGAANAACCTTACCTACCCCTT 978
      961 GACATGTACGAAATTTTCTAGAGATAGATTAGTG-CTTCGGAGAACGCTAACACAGGTGCT 1019
      979 GACATGTACGAAATTTTCTAGAGATAGATTAGTG-CTTCGGAGAACGCTAACACAGGTGCT 1038
      1020 GCATGGCTGTGCTCAGCTGTGTGCTGAGATGTTGGGTTAAATCCCCGCAACAGCGCAAC 1079
      1039 GCATGGCTGTGCTCAGCTGTGTGCTGAGATGTTGGGTTAAATCCCCGCAACAGCGCAAC 1098
      1080 CCTTGCTAATTAATGTCATCATTTGGTTGGGACCTTTAATGAGTACCGCGTGAACAAC 1139
      1099 CCTTGCTAATTAATGTCATCATTTGGTTGGGACCTTTAATGAGTACCGCGTGAACAAC 1158
      1140 GGAGAAAGGTGGAGATGACGTAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACGTA 1199
      1159 GGAGAAAGGTGGAGATGACGTAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACGTA 1218
      1200 ATACAAATGGCGGCTACAGAGGTTGCGCAACCCGCGAGGGGAGCTAATTCGAAAGCGC 1259
      1219 ATACAAATGGCGGCTACAGAGGTTGCGCAACCCGCGAGGGGAGCTAATTCGAAAGCGC 1278
      1260 GTGCTAGTCCGATGCGAGTGTGCAACTGACTCCGCTGAAGTCCGAAATGCTAGTAATCG 1319
      1279 GTGCTAGTCCGATGCGAGTGTGCAACTGACTCCGCTGAAGTCCGAAATGCTAGTAATCG 1338
      1320 CGGATCAGCATGTCCGGGTGAATACGTTCCCGGGCTTTGTACACACCGCGCTCACACA 1379
      1339 CGGATCAGCATGTCCGGGTGAATACGTTCCCGGGCTTTGTACACACCGCGCTCACACA 1398
      1380 TGGAGTGGGTTTCAACGAGACAGTACTTAACCGTAAGAGGCGCTTGCACGCTG 1439
      1399 TGGAGTGGGTTTCAACGAGACAGTACTTAACCGTAAGAGGCGCTTGCACGCTG 1458
      1440 AGATTCAATGACTGGGG 1455
      1459 TGGTCAATGACTTTGGG 1474
      DB
```

```
RESULT 13
US-10-659-948A-18
; Sequence 18, Application US/10659948A
; Publication No. US20040101946A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneec, Timothy A
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-294309
; CURRENT APPLICATION NUMBER: US/10/659,948A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: N. Aestuarii-like AOB P4clone42 16S rDNA
US-10-659-948A-18
```

```
Query Match      93.2%; Score 1357.2; DB 7; Length 1467;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1413; Conservative 0; Mismatches 38; Indels 3; Gaps 3;
```



QY 1 ATTGAACGCTGGCGCATGCTTTACACATGCAAGTCGACCGACGACGATGCTTCAT 60  
DB 16 ATTGAACGCTGGCGCATGCTTTACACATGCAAGTCGACCGACGACGATGCTTCAC 75  
QY 61 CTGTGGCGAGTGGCGGACGGGTGATGATTCGGAACGTAATCCGAAGAAGGGGGTA 120  
DB 76 CTGTGGCGAGTGGCGGACGGGTGATGATTCGGAACGTAATCCGAAGAAGGGGGTA 135  
QY 121 ACCGATCGAAAGATGTGCTAAATACCGCATATACCTAAGAGAGAAAGCAGGGATCGAAA 180  
DB 136 ACCGATCGAAAGATGTGCTAAATACCGCATATACCTAAGAGAGAAAGCAGGGATCGAAA 195  
QY 181 GACCTTCGCTTTTGGAGCGGCGCATGCTGATTAGCTAGTTGGTGGGTAAAGCCTTAC 240  
DB 196 GACCTTCGCTTTTGGAGCGGCGCATGCTGATTAGCTAGTTGGTGGGTAAAGCCTTAC 255  
QY 241 CAAGCGACATCAGTAGTGTGCTGAGAGACGACCGACCACTGGGAATGAGACACG 300  
DB 256 CAAGCGACATCAGTAGTGTGCTGAGAGACGACCGACCACTGGGAATGAGACACG 315  
QY 301 GCCGAGCTCCTTACGGGAGGACGACGATGGGAAATTTTGGACATGGGCGCAAGCCTGATC 360  
DB 316 GCCGAGCTCCTTACGGGAGGACGACGATGGGAAATTTTGGACATGGGCGCAAGCCTGATC 375  
QY 361 CAGCAATGCCCGCTGAGTGAAGAAAGCCTTCGGTGTGAAAGCTTTCACTGAGAAAG 420  
DB 376 CAGCAATGCCCGCTGAGTGAAGAAAGCCTTCGGTGTGAAAGCTTTCACTGAGAAAG 434  
QY 421 AAAGGTTACGCTAAATTAATCTGACCCATGACGCTATGACAGAAAGACCGGCTTAC 480  
DB 435 AAAGGTTACGCTAAATTAATCACTATGATGATGATGATGATGATGATGATGATGATG 494  
QY 481 TACGTCGACGACGCGCGGTAAATCTGATGGGTGCAAGGCTTAATCGAATTAATGAGGCT 540  
DB 495 TACGTCGACGACGCGCGGTAAATCTGATGGGTGCAAGGCTTAATCGAATTAATGAGGCT 554  
QY 541 AAAGGTCGCGAGCGCGCTTGTAAAGTGAATGTGAAATCCCGGCTTAACTGGGAAT 600  
DB 555 AAAGGTCGCGAGCGCGCTTGTAAAGTGAATGTGAAATCCCGGCTTAACTGGGAAT 614  
QY 601 TCGCTTGAATCTACAAAGCTAAGTAGTGGCGACAGGAGGTGGAATTCATGTGTAGCAG 660  
DB 615 TCGCTTGAATCTACAAAGCTAAGTAGTGGCGACAGGAGGTGGAATTCATGTGTAGCAG 674  
QY 661 TGAATGTGCTAGAGTATGGAAGACATCGATGGGAAAGGCGCTCCGGGTAAACT 720  
DB 675 TGAATGTGCTAGAGTATGGAAGACATCGATGGGAAAGGCGCTCCGGGTAAACT 734  
QY 721 GACGCTCAGTCAAGAAAGCTGTGGGAGCAACAGGATTAAGATACCTGTAGTCCAGCC 780  
DB 735 GACGCTCAGTCAAGAAAGCTGTGGGAGCAACAGGATTAAGATACCTGTAGTCCAGCC 794  
QY 781 CTAACGATGTCACTAGTTGTGGGCTTATTAGGCTTGTGTAAGAAAGCTAACCGCTGA 840  
DB 795 CTAACGATGTCACTAGTTGTGGGCTTATTAGGCTTGTGTAAGAAAGCTAACCGCTGA 854  
QY 841 AGTTAACCGCTGTGGGAGTACGCTGTGCAAGATTAACCTCAAGAAAGTTGACGGGAGCC 900  
DB 855 AGTTAACCGCTGTGGGAGTACGCTGTGCAAGATTAACCTCAAGAAAGTTGACGGGAGCC 914  
QY 901 GCACAAGCGGTGATTAATGATGATTAATCGATGCAACGCAAAACCTTAACCTTACCTT 960  
DB 915 GCACAAGCGGTGATTAATGATGATTAATCGATGCAACGCAAAACCTTAACCTTACCTT 974  
QY 961 GACATGTAGCGAATTTCTAGAGATGATTAAGTGC-TTCGGAAAGCTTAACAGAGTGTCT 1019  
DB 975 GACATGTAGCGAATTTCTAGAGATGATTAAGTGC-TTCGGAAAGCTTAACAGAGTGTCT 1034  
QY 1020 GCATGGCTGTGCTGAGCTGTGTGCTGAGATGTTGGGTAAATCCCGCAACAGAGGCAAC 1079  
DB 1035 GCATGGCTGTGCTGAGCTGTGTGCTGAGATGTTGGGTAAATCCCGCAACAGAGGCAAC 1094

QY 1080 CCTGTCAATTAATGGCATCATTTGGTGGGCACTTTAATGAGACTGCGGTGACAAACC 1139  
DB 1095 CCTGTCAATTAATGGCATCATTTAGTGGGCACTTTAATGAGACTGCGGTGACAAACC 1154  
QY 1140 GGAGAAAGGTGGGAGTGAAGTCAAGTCTTCATGAGCCCTTAATGGGTAGGCTTACACGTA 1199  
DB 1155 GGAGAAAGGTGGGAGTGAAGTCAAGTCTTCATGAGCCCTTAATGGGTAGGCTTACACGTA 1214  
QY 1200 ATACATGCGCGGTACAGAGGTTGGCAACCCCGGAGGGGAGACTAATCTCAGAAAGCGC 1259  
DB 1215 ATACATGCGCGGTACAGAGGTTGGCAACCCCGGAGGGGAGACTAATCTCAGAAAGCGC 1274  
QY 1260 GTCTGATCCGATTCGAGTCTGCAACTCGACTCCGTGAAGTGCAGATCGCTAGTATCG 1319  
DB 1275 GTCTGATCCGATTCGAGTCTGCAACTCGACTCCGTGAAGTGCAGATCGCTAGTATCG 1334  
QY 1320 CGGATCAGCATGTGCGGTGAATACGTTCCGGGCTCTTGTACACACCGCCGTACACCA 1379  
DB 1335 CGGATCAGCATGTGCGGTGAATACGTTCCGGGCTCTTGTACACACCGCCGTACACCA 1394  
QY 1380 TGGAGTGGGTTTACACAGAGCAGTATGCTTAACCTTAAGAGGCGGCTTGCACGCTG 1439  
DB 1395 TGGAGTGGGTTTACACAGAGCAGTATGCTTAACCTTAAGAGGCGGCTTGCACGCTG 1453  
QY 1440 AGATTCAATGACTGG 1453  
DB 1454 AGATTCAATGACTGG 1467

RESULT 14  
US-10-659-980a-18  
; Sequence 18, Application US/10659980A  
; Publication No. US20040106135A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovaneq, Timothy A  
; TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-284781  
; CURRENT FILING DATE: US/10/659, 980A  
; PRIOR APPLICATION NUMBER: US 2003-09-10  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386, 217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386, 218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386, 219  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 1467  
; TYPE: DNA  
; ORGANISM: Unknown  
; OTHER INFORMATION: N. Aestuarii-like AOB pAc1one42 16S rDNA  
US-10-659-980a-18

Query Match 93.2%; Score 1357.2; DB 7; Length 1467;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1413; Conservative 0; Mismatches 38; Indels 3; Gaps 3;

QY 1 ATTGAACGCTGGCGCATGCTTTACACATGCAAGTCGACCGACGACGATGCTTCAT 60  
DB 16 ATTGAACGCTGGCGCATGCTTTACACATGCAAGTCGACCGACGACGATGCTTCAC 75  
QY 61 CTGTGGCGAGTGGCGGACGGGTGATGATTCGGAACGTAATCCGAAGAAGGGGGTA 120  
DB 76 CTGTGGCGAGTGGCGGACGGGTGATGATTCGGAACGTAATCCGAAGAAGGGGGTA 135  
QY 121 ACCGATCGAAAGATGTGCTAAATACCGCATATACCTAAGAGAGAAAGCAGGGATCGAAA 180  
DB 136 ACCGATCGAAAGATGTGCTAAATACCGCATATACCTAAGAGAGAAAGCAGGGATCGAAA 195



Db 376 CAGCAATGCCCGTGAAGTGAAGAGG-CCTCGGGTTTAAAGCTCTTTCAGTGAAGAGA 434  
Qy 421 AAAGTTACGGTAATTAATCGTGAACCCATGACGGTATCGA CAGAAGAGACCGGGCTAAC 480  
Db 435 AAAGTTGTACATTAATTAATCAACAATTATGTGTATCGA CAGAAGAGACCGGGCTAAC 494  
Qy 481 TACGTGCGACAGCCCGGGTAAATCGTAGGGTGCAGACCGTTAATCGAATTAATCTGGGGCT 540  
Db 495 TACGTGCGACAGCCCGGGTAAATCGTAGGGTGCAGACCGTTAATCGAATTAATCTGGGGCT 554  
Qy 541 AAAGGGGCGCAGCGCGGCTTGTAAAGCAGATGGAATCCCGGGCTTAACCTGGGAAT 600  
Db 555 AAAGGGGCGCAGCGCGGCTTGTAAAGCAGATGGAATCCCGGGCTTAACCTGGGAAT 614  
Qy 601 TGGCTTTGAATCAACAAGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTACAG 660  
Db 615 TGGCTTTGAATCAACAAGCTAGAGTGTGCGAGAGGAGGAGGTGAATTCATGTGTACAG 674  
Qy 661 TGAATGCGTAGAGATATGGAAGAACATCGATGCGAAGGACGCTCTGGGTAACT 720  
Db 675 TGAATGCGTAGAGATATGGAAGAACATCGATGCGAAGGACGCTCTGGGTAACT 734  
Qy 721 GACGCTATGCAACGAAGCGTGGGAGCAACAGGATTAAGTACCTGGTATGTCAGAGC 780  
Db 735 GACGCTATGCAACGAAGCGTGGGAGCAACAGGATTAAGTACCTGGTATGTCAGAGC 794  
Qy 781 CTAAACAGATGCACTAGTTGTTGGGCTTATTAAGCTGTGTAACGAAGCTAACGCGTGA 840  
Db 795 CTAAACAGATGCACTAGTTGTTGGGCTTATTAAGCTGTGTAACGAAGCTAACGCGTGA 854  
Qy 841 AGTTGACCGCTGGGAGAGTACGCTGCGAAGTTAAACTCAAGAAATGACGGGAGCC 900  
Db 855 AGTTGACCGCTGGGAGAGTACGCTGCGAAGTTAAACTCAAGAAATGACGGGAGCC 914  
Qy 901 GCACAACCGGTGATTAATGTGGATTAATTCGATGCAACGCGAAGACCTTAACCTACCTT 960  
Db 915 GCACAACCGGTGATTAATGTGGATTAATTCGATGCAACGCGAAGACCTTAACCTACCTT 974  
Qy 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC-TTCGGGAACGCTTAACACAGTGTCT 1019  
Db 975 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC-TTCGGGAACGCTTAACACAGTGTCT 1034  
Qy 1020 GCATGGCTGTGTCAGTCTGTGTCGAGATGTTGGGTTAAATGCCGCAACGACGCGAAC 1079  
Db 1035 GCATGGCTGTGTCAGTCTGTGTCGAGATGTTGGGTTAAATGCCGCAACGACGCGAAC 1094  
Qy 1080 CCTTGTCAATTAATGTCCATCAATTTGGTGGGCACTTTAATGAGACTGCCGCTGACAAAC 1139  
Db 1095 CCTTGTCAATTAATGTCCATCAATTTGGTGGGCACTTTAATGAGACTGCCGCTGACAAAC 1154  
Qy 1140 GGAGGAAGGTGGGATGAGCTCAAGTCTCATGGCCCTTATGGGTAGGGCTTCAACAGTA 1189  
Db 1155 GGAGGAAGGTGGGATGAGCTCAAGTCTCATGGCCCTTATGGGTAGGGCTTCAACAGTA 1214  
Qy 1200 ATACAAATGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGGCTAATCTCAGAAAGGCGC 1259  
Db 1215 ATACAAATGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGGCTAATCTCAGAAAGGCGC 1274  
Qy 1260 GTCTGTAGTCCGGATCGAGTCTGCAACTCGACTCCGTGAAGTCCGAATCGCTAGTAATCG 1319  
Db 1275 GTCTGTAGTCCGGATCGAGTCTGCAACTCGACTCCGTGAAGTCCGAATCGCTAGTAATCG 1334  
Qy 1320 CGGATTCAGCATGTGCGCGGTGAATAGCTTCCGGGGCTTGTACACACGCGCGCTACACCA 1379  
Db 1335 CGGATTCAGCATGTGCGCGGTGAATAGCTTCCGGGGCTTGTACACACGCGCGCTACACCA 1394  
Qy 1380 TGGGAGTGGGTTTCAACAGAAAGAGTGTAAACCGTAAAGAGGCGCTTGGCAACGGTGT 1439  
Db 1395 TGGGAGTGGGTTTCAACAGAAAGAGTGTAAACCGTAAACCGTAAACCGTAAACCGTGT 1453  
Qy 1440 AGATTCAATGACTGG 1453

Db 1454 AGATTCAATGACTGG 1467  
RESULT 16  
US-10-659-948A-3  
; Sequence 3, Application US/10659948A  
; Publication No. US20040101946A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovaneec, Timothy A  
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-294309  
; CURRENT APPLICATION NUMBER: US/10/659,948A  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573,684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386,217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,219  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 1458  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: AOB Type B R3clones 16S rDNA  
US-10-659-948A-3

Query Match 88.0%; Score 1282.8; DB 7; Length 1458;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 1363; Conservative 0; Mismatches 92; Indels 5; Gaps 2;  
Qy 1 ATGGAACGCTGGGCGCATGCTTACACATGCAAGTCGAACGCGACGACGATGCTTGAT 60  
Db 1 ATTGAACGCTGGGCGCATGCTTACACATGCAAGTCGAACGCGACGACGCGGGGAAC-- 58  
Qy 61 CTGGTGGCGAGTGGCGGACGGGTGATGATCATTCGGAACGTATCCGAAGAGGGGGGATA 120  
Db 59 CTGGTGGCGAGTGGCGGACGGGTGATGATCATTCGGAACGTATCCGAAGAGGGGGGATA 118  
Qy 121 ACCGATCGAAGATGTGCTTAATCCGATTAATCTTAAGGAGGAAAGCAGGGGATCGAAA 180  
Db 119 ACCGATCGAAGATGTGCTTAATCCGATTAATCTTAAGGAGGAAAGCAGGGGATCGAAA 178  
Qy 181 GACCTTGCGCTTTTGGAGCGCGCATGTCTGATTAGCTAGTGTGGGTAAAGGCTTAC 240  
Db 179 GACCTTGCGCTTTTGGAGCGCGCATGTCTGATTAGCTAGTGTGGGTAAAGGCTTAC 238  
Qy 241 CAAGCGACGATGATGTTGTCTGAGAGGACGACCGACACTGGGACTGAGACAG 300  
Db 239 CAAGCGACGATGATGTTGTCTGAGAGGACGACCGACACTGGGACTGAGACAG 298  
Qy 301 GCCGAGCTCTTAACGCGAGGACGAGTGGGAAATTTGGCAATGGCGCAAGCTTATC 360  
Db 299 GCCGAGCTCTTAACGCGAGGACGAGTGGGAAATTTGGCAATGGCGGAAACCTTGATC 358  
Qy 361 CAGCAATGCGCGGTGATGGAAGAGCGCTTGGGTTGAAGCTCTTTCAGTGCAGAGA 420  
Db 359 CAGCAATGCGCGGTGATGGAAGAGCGCTTGGGTTGAAGCTCTTTCAGTGCAGAGA 418  
Qy 421 AAAGTTACCGTAAATTAATCGTGAACCCATGACGGTATCGA CAGAAGAGACCGGGCTAAC 480  
Db 419 AAAGTTACCGTAAATTAATCGTGAACCCATGACGGTATCGA CAGAAGAGACCGGGCTAAC 478  
Qy 481 TACGTGCGACAGCCCGGGTAAATCGTAGGGTGCAGACCGTTAATCGAATTAATCTGGGGCT 540  
Db 479 TACGTGCGACAGCCCGGGTAAATCGTAGGGTGCAGACCGTTAATCGAATTAATCTGGGGCT 538  
Qy 541 AAAGGGGCGCAGCGCGGCTTGTAAAGCAGATGGAATCCCGGGCTTAACCTGGGAAT 600

```
Db 539 AAAGCTGCGCAGCGCGTTTGTAACTCAGATGAAAGCCCGGGCTTAACCTGGGAAC 598
Qy 601 TGGTTTGAACCTAACAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGACAG 660
Db 599 TGGTTTGAACCTAACAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCACAGGTAGCAG 658
Qy 661 TGAATGCGTAGAGATATGGAAGAACATCGATGCGAGACACCTCCGTGGTTAACTACT 720
Db 659 TGAATGCGTAGAGATATGGAAGAACACCGATGGCGAAGGACCCCTGGTTAAACACC 718
Qy 721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCCAGCC 780
Db 719 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCCAGCC 778
Qy 781 CTAAACGATGTCAACTAGTTGTTGGCCCTTATTAGGCTGTGTAACGAAGCTAACCGCTGA 840
Db 779 CTAAACGATGTCAACTAGTTGTTGGCCCTTATTAGGCTGTGTAACCGACTTAACGGCTGA 838
Qy 841 AGTTGACCGCTGGGAGTACCGTCCGAAAGATTAAATCTCAAGAAATTGACGGGGACCC 900
Db 839 AGTTGACCGCTGGGAGTACCGTCCGAAAGATTAAATCTCAAGAAATTGACGGGGACCC 898
Qy 901 GCAACAGCGGTGGAATTAATGATGATTAATTGATGCAACGCGAATAAATCTTACCTACCTT 960
Db 899 GCAACAGCGGTGGAATTAATGATGATTAATTGATGCAACGCGAATAAATCTTACCTACCTT 958
Qy 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTC---TTGGGAAACGCTAACACAGGTG 1017
Db 959 GACATGTAGCGAAGCCCGCGAGAGGTGGTGTGCCGAAAGGAGCGGTAAACACAGGTG 1018
Qy 1018 CTGCAATGCGTGTGTCACTGCTGTGTGTGATGATGTTGGGTAAAGTCCCGCAAGACGCA 1077
Db 1019 CTGCAATGCGTGTGTCACTGCTGTGTGTGATGATGTTGGGTAAAGTCCCGCAAGACGCA 1078
Qy 1078 ACCCTGTCACTAATTGCGCATCACTTGTGGGCACTTTAATGAGACTGCGGTGACAA 1137
Db 1079 ACCCTGTCACTAATTGCGCATCACTTGTGGGCACTTTAATGAGACTGCGGTGACAA 1138
Qy 1138 CCGAGAGAGGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGGGTAGGGCTTCAACG 1197
Db 1139 CCGAGAGAGGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGGGTAGGGCTTCAACG 1198
Qy 1198 TAATCAATGCGCGCTGACAGAGGTTGCCAACCCCGAGGGAGAGCTAACTGAGAAAGC 1257
Db 1199 TAATCAATGCGCGCTGACAGAGGTTGCCAACCCCGAGGGAGAGCTAACTGAGAAAGC 1258
Qy 1258 GCGTCGATGCGGAGTGGAGTCTGCAACTCGACTCGGTGAAGTGGGAATCGCTAGTAAT 1317
Db 1259 GCGTCGATGCGGAGTGGAGTCTGCAACTCGACTCGGTGAAGTGGGAATCGCTAGTAAT 1318
Qy 1318 CCGGATCAGCATGTGCGGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCAAC 1377
Db 1319 CCGGATCAGCATGTGCGGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCAAC 1378
Qy 1378 CATGGAGTGGGTTTCAACAGAGCAGTAGTCTTAACCGTAAGAGGGCGCTTGCAACGG 1437
Db 1379 CATGGAGTGGGTTTCAACAGAGCAGTAGTCTTAACCGTAAGAGGGCGCTTGCAACGG 1438
Qy 1438 TGAGATTCACTGAGGCTG 1457
Db 1439 TGAGATTCACTGAGGCTG 1458
```

## RESULT 17

```
US-10-659-980A-3
; Sequence 3, Application US/10659980A
; Publication No. US20040106133A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneec, Timothy A
; TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284781
; CURRENT APPLICATION NUMBER: US/10/659,980A
; CURRENT FILING DATE: 2003-09-10
```

```
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1458
; TYPE: DNA
; ORGANISM: Unknown
; OTHER INFORMATION: AOB Type B R3clones 16S rDNA
US-10-659-980A-3
```

```
Query Match 88.0%; Score 1282.8; DB 7; Length 1458;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1363; Conservative 0; Mismatches 92; Indels 5; Gaps 2;
```

```
Qy 1 ATTGAACGCTGGCGGCATGCTTTACATGCAATGCAACGCGACACGAGTCTTCAT 60
Db 1 ATTGAACGCTGGCGGCATGCTTTACATGCAATGCAACGCGACACGCGGGCAAC-- 58
Qy 61 CTGGTGGCAGTGGGAGCGGCTGATATGATGCGAAGCTATCAGAAAGAGGGGGGTA 120
Db 59 CTGGTGGCAGTGGGAGCGGCTGATATGATGCGAAGCTATCAGAAAGAGGGGGGATTA 118
Qy 121 ACGCATCGAAAGATGTGCTAATACCGCATATCTCTAAGAGAAAGCAGGGGATCGAA 180
Db 119 ACGCATCGAAAGATGTGCTAATACCGCATATCTCTAAGAGAAAGCAGGGGATCGAA 178
Qy 181 GACCTTGCGCTTTTGAAGCGCGCATGTCTGATTAGTATTGTTGGGTAAAGCCTTAC 240
Db 179 GACCTTGCGCTTTTGAAGCGCGCATGTCTGATTAGTATTGTTGGGTAAAGCCTTAC 238
Qy 241 CAAAGCGAGATCAGTATGTTGTTGAGAGGACGACACGCGACCTGGGACTGAGCACG 300
Db 239 CAAAGCGAGATCAGTATGTTGTTGAGAGGACGACACGCGACCTGGGACTGAGCACG 298
Qy 301 GCCCAAGCTCTTACGGGAGGAGCAGAGTGGGGAATTTTGGACATAGGGCGCAAGCTGATC 360
Db 299 GCCCAAGCTCTTACGGGAGGAGCAGAGTGGGGAATTTTGGACATAGGGGGAACCTGATC 358
Qy 361 CAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTTGAAGCTCTTTCACTGAGAGAA 420
Db 359 CAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTTGAAGCTCTTTCACTGAGAGAA 418
Qy 421 AAAGTTAAGGTAAATTAATCGTGAACCCATGACGGTATGACAGAAAGACCGGCTAAC 480
Db 419 AAAGTTAAGGTAAATTAATCGTGAACCCATGACGGTATGACAGAAAGACCGGCTAAC 478
Qy 481 TAGTCCAGACCGCGGTAAATAGCTAGGTTGCAAGCCTTATCGAATTAATCTGGGCGT 540
Db 479 TAGTCCAGACCGCGGTAAATAGCTAGGTTGCAAGCCTTATCGAATTAATCTGGGCGT 538
Qy 541 AAAGGTGGCAGGCGGCTTGTAACTCAGATGTGAATCCCGGGCTTAACCTGGGAAT 600
Db 539 AAAGGTGGCAGGCGGCTTGTAACTCAGATGTGAATCCCGGGCTTAACTGGGAAC 598
Qy 601 TGGTTTGAACCTAACAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGACAG 660
Db 599 TGGTTTGAACCTAACAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCACAGGTAGCAG 658
Qy 661 TGAATGCGTAGAGATATGGAAGAACATCGATGCGAGACACCTCCGTGGTTAACTACT 720
Db 659 TGAATGCGTAGAGATATGGAAGAACACCGATGGCGAAGGACCCCTGGTTAAACACC 718
Qy 721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCCAGCC 780
Db 719 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCCAGCC 778
```

Qy	781	CTAAACGATGCACTAGTTGTTGGGACCTTATTTAGGCTTGCTTAACGAAGCTAACGCGCTGA	840
Db	779	CTAAACGATGCACTAGTTGTTGGGACCTTATTTAGGCTTGCTTAACGAAGCTTAACGCGCTGA	838
Qy	841	AGTTGACCCGCTTGGGGAGTACGGTTCGACAGATTAAATCTCAAGAAATTGACGGGGACCC	900
Db	839	AGTTGGCCGCTGGGGAGTACGGTTCGACAGATTAAATCTCAAGAAATTGACGGGGACCC	898
Qy	901	GCACAAAGGGTGGATTAATGTGGATTAAATTCGATCGCAACGGCAAAACCTTAACCTT	960
Db	899	GCACAAAGGGTGGATTAATGTGGATTAAATTCGATCGCAACGGCAAAACCTTAACCTT	958
Qy	961	GACATGTGACGAATTTCTAGAGATAGATTAGTGC---TTGGGGAAGGCTTAACAGGTG	1017
Db	959	GACATGTACCGAAGCCCGCCGAGAGTGGGTGTGCTCCGAAAGGAGGGGTAAACAAGGTG	1018
Qy	1018	CTGCATGGCTGTCTGACGTCGTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACAGCGCA	1077
Db	1019	CTGCATGGCTGTCTGACGTCGTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACAGCGCA	1078
Qy	1078	ACCTTGTGATTTAATGGCATTCTTGGTTGGCACTTATGAGACTGCGCGTGACAA	1137
Db	1079	ACCTTGTGATTTAATGGCATTCTTGGTTGGCACTTATGAGAACTGCGGTGACAA	1138
Qy	1138	CCGAGAGAAAGGGGGATGACGTCAAGTCTCATGACCCTTATGAGGTAGGGCTTACACG	1197
Db	1139	CCGAGAGAAAGGGGGATGACGTCAAGTCTCATAGCCCTTATGAGGTAGGGCTTACACG	1198
Qy	1198	TAAATCAATGGCGCGCTACAGAGAGGTTGCCAACCCGCGAGGGGGAGCTAATCTCAGAAAGC	1257
Db	1199	TAAATCAATGGCGCGCTACAGAGAGGTTGCCAACCCGCGAGGGGGAGCTAATCTCAGAAAGC	1258
Qy	1258	GGGTGTGATCCGGAATCCGAGTCTGCCAATCGAATCGGTGAAGTCCGAATCGTAT	1317
Db	1259	GGGTGTGATCCGGAATCCGAGTCTGCCAATCGAATCGGTGAAGTCCGAATCGTAT	1318
Qy	1318	CGCGATCAGCATGTCCGGGTGAAATACGTTCCGGGTCTTGTACACACCGCCGTCACAC	1377
Db	1319	CGCGATCAGCATGTCCGGGTGAAATACGTTCCGGGTCTTGTACACACCGCCGTCACAC	1378
Qy	1378	CATGGAGTGGGTTTACCCAGAGCAGGTAGTCTAACCGTAAAGAGGGCGCTTGCACAG	1437
Db	1379	CATGGAGTGGGTTTACCCAGAGCAGGTAGTCTAACCGCAGAGGGCGCTTGCACAG	1438
Qy	1438	TGAGATTCAATGACTGGGGTG 1457	
Db	1439	TGAGATTCAATGACTGGGGTG 1458	
RESULT 18			
US-10-659-983A-3			
/ Sequence 3, Application US/10659983A			
/ Publication No. US2004015731A1			
/ GENERAL INFORMATION:			
/ APPLICANT: Hovaneec, Timothy A			
/ TITLE OF INVENTION: Ammonia-Oxidizing Bacteria			
/ FILE REFERENCE: 81289-284779			
/ CURRENT APPLICATION NUMBER: US/10/659,983A			
/ CURRENT FILING DATE: 2003-09-10			
/ PRIOR APPLICATION NUMBER: US-09/573,684			
/ PRIOR FILING DATE: 2000-05-19			
/ PRIOR APPLICATION NUMBER: US 60/386,217			
/ PRIOR FILING DATE: 2002-09-19			
/ PRIOR APPLICATION NUMBER: US 60/386,218			
/ PRIOR FILING DATE: 2002-09-19			
/ PRIOR APPLICATION NUMBER: US 60/386,219			
/ PRIOR FILING DATE: 2002-09-19			
/ NUMBER OF SEQ ID NOS: 23			
/ SOFTWARE: PatentIn version 3.2			
/ SEQ ID NO 3			
/ LENGTH: 1458			
/ TYPE: DNA			

Query Match	Best Local Similarity	Score	DB	Length
88.0%;	93.4%;	1282.8;	7;	1458;
Matches 1363;	Conservative 0;	Mismatches 92;	Indels 5;	Gaps 2.

  

Query	Match	Similarity	Score	DB	Length
1	ATTGACGCTGCGCGCATGCTTTACATCATGGAAGTCGAAAGCGGACGACGATGCTTGCAT	60			
Db	1	ATTGAAAGCTTGGCGGACGCTTTTACATCATGGAAGTCGAAAGCGGACGACGCGGCGAAC--	58		
Qy	61	CTGGTGGCGAGTGGCGGACGCGGTAGTATCATCTGAAACGTATTCAGAAAGAGGGGGGTA	120		
Db	59	CTGGTGGCGAGTGGCGGAAACGGGTAGTATCATCTGAAACGTATTCAGAAAGAGGGGGGTA	118		
Qy	121	ACGCATCGAAAGATGTGCTAAATACCGCATATCTTAAAGAGGAAACGAGGGGATCGAAA	180		
Db	119	ACGCACCGAAAGTGTGCTAAATACCGCATATCTTCAAGAGGAAAGAGGGGATCGCAA	178		
Qy	181	GACCTTGGGCTTTTGGAGCGGCGCATGCTATTAAGTATGCTGGTGGGGTTAAAGGCTTAC	240		
Db	179	GACCTTGGGCTTTTGGAGCGGCGCATGCTATTAAGTATGCTGGTGGGGTTAAAGGCTTAC	238		
Qy	241	CAAGGCGACGATCATAGTATGCTGAGAGGACGACGACCACTGGGACTGAGACAG	300		
Db	239	CAAGGCGACGATCATAGTATGCTGAGAGGACGACGACCACTGGGACTGAGACAG	298		
Qy	301	GCCGACACTCTTACCGGAGGCGACAGTGGGAAATTTGGACATGGGCGCAAGCCTGATC	360		
Db	299	GCCGACACTCTTACCGGAGGCGACAGTGGGAAATTTGGACATGGGCGGAAACCTGATC	358		
Qy	361	CAGCATGCTCCGCTGAGAGGAAAGGCGGCTTGAAGCTCTTTCAGTCCGAGAA	420		
Db	359	CAGCATGCTCCGCTGAGAGGAAAGGCGGCTTGAAGCTCTTTCAGTCCGAGAA	418		
Qy	421	AAAGGTTACGGTAAATTAATCGTGAACCCATGACGATGACGACGAAAGAACCCGCTAAC	480		
Db	419	AAAGGTTACGGTAAATTAATCGTGAACCCATGACGATGACGACGAAAGAACCCGCTAAC	478		
Qy	481	TACGTGCCAGCAGCCGCGGTAAATCGTAGGCTGCAAGCGTTAATGGAATTACTGGGCT	540		
Db	479	TACGTGCCAGCAGCCGCGGTAAATCGTAGGCTGCAAGCGTTAATGGAATTACTGGGCT	538		
Qy	541	AAAGGTGGCGAGCGGCGCTTGTAGTCAAGTGAATATCCCGGGCTTAACTGGGAAT	600		
Db	539	AAAGGTGGCGAGCGGCGCTTGTAGTCAAGTGAATATCCCGGGCTTAACTGGGAAT	598		
Qy	601	TGCGTTTGAATCTAACAAAGCTAGATGTGGCGAGAGGAGTGGAAATTCATGTGTAGCAG	660		
Db	599	TGCGTTTGAATCTAACAAAGCTAGATGTGGCGAGAGGAGGAGTGGAAATTCATGTGTAGCAG	658		
Qy	661	TGAAATGCGTAGAGTATGGAAGAACATGATGGCGAAGCGCCTCTGGGTTAACT	720		
Db	659	TGAAATGCGTAGAGTATGGAAGAACACCGATGGCGAAGCGCCTCTGGGTTAACT	718		
Qy	721	GACGCTCATGCAAGAAAGCGTGGGAGCAACAGAGTTAGATATCCCTGGAGTCCAGCC	780		
Db	719	GACGCTCATGCAAGAAAGCGTGGGAGCAACAGAGTTAGATATCCCTGGAGTCCAGCC	778		
Qy	781	CTAAACGATGTCACTAGTGTGTGGGCTTTATTAAGGCTTGTATCGAAGCTTAAAGCGTGA	840		
Db	779	CTAAACGATGTCACTAGTGTGTGGGCTTTATTAAGGCTTGTATCGAAGCTTAAAGCGTGA	838		
Qy	841	AGTTGACGCTGGGAGGATAGCGTGCAGATTTAAATCTCAAGAAATTTGACGGGGACCC	900		
Db	839	AGTTGACGCTGGGAGGATAGCGTGCAGATTTAAATCTCAAGAAATTTGACGGGGACCC	898		
Qy	901	GCAAGGCGGTGATTAATGATTTAATTCGATGCAAGCGGAAACCTTACCTACCTT	960		
Db	899	GCAAGGCGGTGATTAATGATTTAATTCGATGCAAGCGGAAACCTTACCTACCTT	958		



Db 1138 AACCCGAGGAGGTGGAGTGAAGTCAAGTCTCAAGGCCCTTATGGGAGGCTTAC 1197  
Qy 1195 ACCATAACAAATGCGCGTACAGAGGTTGCAACCCGAGGGGAGCTAATCTCAGAA 1254  
Db 1198 ACCATAACAAATGCGCGTACAGAGGTTGCAACCCGAGGGGAGGCAATCTCAGAA 1257  
Qy 1255 AGCGCGTCTAGTCCGATCCGAGTCTGCAACTGCACTCCGTGAAGTCGGAATCGTAGT 1314  
Db 1258 AGCAGCTCTAGTCCGATCCGAGTCTGCAACTGCACTCCGTGAAGTCGGAATCGTAGT 1317  
Qy 1315 AATCGCGATCAGCATCTCGCGTGAATACGTTCCGGGCTTTTACACACCCGCGTCA 1374  
Db 1318 AATCGCGATCAGCATCCGCGGTGAATACGTTCCGGGCTTTTACACACCCGCGTCA 1377  
Qy 1375 CACCATGGAGTGGGTTTCCACAGACAGGTAGCTTAACCGTAAGAGGGCGCTTGCA 1434  
Db 1378 CACCATGGAGTGGGTTTCCACAGACAGGTAGCTTAACCGTAAGAGAGCGCTTGCA 1437  
Qy 1435 CGGTGAGATTCACTGAGTGGGTTG 1457  
Db 1438 CGGTGAGGATCATCACTGGGTTG 1460

RESULT 20  
US-10-659-980A-4  
Sequence 4, Application US/10659980A  
Publication No. US20040106133A1  
GENERAL INFORMATION:  
APPLICANT: Hovaneq, Timothy A  
TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria  
FILE REFERENCE: 81289-284781  
CURRENT APPLICATION NUMBER: US/10/659,980A  
CURRENT FILING DATE: 2003-09-10  
PRIOR APPLICATION NUMBER: US 09/573,684  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 60/386,217  
PRIOR FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: US 60/386,218  
PRIOR FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: US 60/386,219  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 4  
LENGTH: 1460  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: AOB Type C R5clone47 16S rDNA  
US-10-659-980A-4

Query Match 84.1%; Score 1225.4; DB 7; Length 1460;  
Best Local Similarity 92.1%; Pred. No. 0;  
Matches 1348; Conservative 0; Mismatches 106; Indels 9; Gaps 5;

Qy 1 ATTGAACGCTGGCGGCAATGCTTTACACATGCAAGTCGAAACCGCACACAGAGTCTGCAT 60  
Db 1 ATTGAACGCTGGCGGCAATGCTTTACACATGCAAGTCGAAACCGCAAC--CGGGGGCTTCGCGC 58  
Qy 61 CTGGTGGCGAGTGGCGGAGCGGGTGAATATGATCGGAACGTAATCCAGAGAGGGGGGTA 120  
Db 59 CTGGCGGCGAGTGGCGGAGCGGGTGAATATGATCGGAACGTAATCTTAAAGTGGGGAATA 118  
Qy 121 ACGCATCGAAGATGTCTAATATCCGCATATCTTAAAGAGAGAAAGCAGGGGATCGAAA 180  
Db 119 ACGCATCGAAGATGTCTAATATCCGCATAT--CTCTGAGAGAGAAAGCAGGGGATCGCAA 177  
Qy 181 GACCTTTGGCTTTTGAAGCGCGGATGCTGTGATTGCTAAGTGGGGGTAAAGGCTTAC 240  
Db 178 GACCTTTGGCTTTAAGAGCGCGGATGCTGTGATTGCTAAGTGGGGGTAAAGGCTTAC 237  
Qy 241 CAAGCGACGATCAGTATGTTGGTCTGAGAGGAGCAGACCACTGGGACTGAGACAG 300

Db 238 CAAGCGACGATCAGTATGTTGGTCTGAGAGAGACCAACCACTGGGACTGAGACAG 297  
Qy 301 GCCCAGACTCTTACGGGAGGCGACAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360  
Db 298 GCCCAGACTCTTACGGGAGGCGACAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 357  
Qy 361 CAGCAATGCCCGCTGAGTGAAGAAGGCTTCGGGTTGTAAAGCTCTTTCAGTGAGAGA 420  
Db 358 CAGCCATGCGCGCTGAGTGAAGAAGGCTTCGGGTTGTAAAGCTCTTTCAGTGAAGA 417  
Qy 421 AAAGTTACGGTAAATATCTGTATCCATGACCGGTATCCGACAGAAAGCAGCGGCTAAC 480  
Db 418 AAAGATCATGAT 477  
Qy 481 TAGGTGCAAGCAGCCCGGTATATAGTGGGTCAAGCGTTAATCCGATTAATCTGGGCGT 540  
Db 478 TAGGTGCAAGCAGCCCGGTATATAGTGGGTCAAGCGTTAATCCGATTAATCTGGGCGT 537  
Qy 541 AAAGGTGCGCAGCGCGCTTGTAAATGTAATGTAATCCCGGGCTTAACCTGGGAAAT 600  
Db 538 AAAGGTGCGCAGCGCGCTTGTAAATGTAATGTAATCCCGGGCTTAACCTGGGAAAT 597  
Qy 601 TGGGTTGAACCTAACAAGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTACAG 660  
Db 598 TGGGTTGAACCTAACAAGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTACAG 657  
Qy 661 TGAATATCGTAGATATGAGAGAACATCGATGCGAAGGCGACCTCTGGGTTAACACT 720  
Db 658 TGAATATCGTAGATATGAGAGAACATCGATGCGAAGGCGACCTCTGGGTTAACACT 717  
Qy 721 GAGCTCATGACGAAAGCGTGGGAGCAACAGATTTAATACCTGGTATGTCACAGCC 780  
Db 718 GAGCTCATGACGAAAGCGTGGGAGCAACAGATTTAATACCTGGTATGTCACAGCC 777  
Qy 781 CTAAACGATGCAACTGTTGTTGGGCTTATTA--GGCTTGTATGCAAGCTAACCGT 838  
Db 778 CTAAACGATGCAACTGTTGTTGGGCTTATTA--GGCTTGTATGCAAGCTAACCGT 837  
Qy 839 GAACTTACCGCTGGGAGATACGATGCGCAAGATTTAAACTCAAGGAATTTGACGGGAGC 898  
Db 838 GAACTTACCGCTGGGAGATACGATGCGCAAGATTTAAACTCAAGGAATTTGACGGGAGC 897  
Qy 899 CCGCACAAGGGGTGATTAATGATTAATTCATGCAAGCGGAAACCTTACCTTACC 958  
Db 898 CCGCACAAGGGGTGATTAATGATTAATTCATGCAAGCGGAAACCTTACCTTACC 957  
Qy 959 TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTGC--TTGCGGAACCTTAACAGAG 1015  
Db 958 TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTGC--TTGCGGAACCTTAACAGAG 1017  
Qy 1016 TGCTGCATGAGCTGCTGACGCTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG 1075  
Db 1018 TGCTGCATGAGCTGCTGACGCTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG 1077  
Qy 1076 CAACCTTGTCAATTAATTTGCCATCAT--TGGTGGGCACTTTAATGAGACTGCGGGTAC 1134  
Db 1078 CAACCTTGTCAATTAATTTGCCATCAT--TGGTGGGCACTTTAATGAGACTGCGGGTAC 1137  
Qy 1135 AAACCGAGAGAAAGTGGGAGTGAAGTCAAGTCCATAGGCCCTTAATGGGTGGGCTTAC 1194  
Db 1138 AAACCGAGAGAAAGTGGGAGTGAAGTCCATAGGCCCTTAATGGGTGGGCTTAC 1197  
Qy 1195 ACGTAAATCAATGCGCGTACAGAGGGTTGCCAACCCGCGAGGGGAGCAATCTCAGAA 1254  
Db 1198 ACGTAAATCAATGCGCGTACAGAGGGTTGCCAACCCGCGAGGGGAGCAATCTCAGAA 1257  
Qy 1255 AGCGGTCTGATGTCGAGTCCGAGTCTGCAACTGCACTCCGTGAAGTGGGAAATGGCTAGT 1314  
Db 1258 AGCAGTCTGATGTCGAGTCCGAGTCTGCAACTGCACTCCGTGAAGTGGGAAATGGCTAGT 1317  
Qy 1315 AATCGCGATCAGATGTCGGGAGTAATGAGTCCCGGGCTTGTACACACCGCGCTCA 1374



Db 1318 AATCGGGATCAGATGCGCGGGTGAATACGTTCCGGGTCTGTATACACACGCGCGTCA 1377  
QY 1375 CACCATGGAGTGGGTTTCCAGAAAGCAGTACTTAACCGTAAAGAGGGCGCTTGCCA 1434  
Db 1378 CACCATGGAGTGGTGTTCACAGAAAGCAGTAGTTAAACGTAAAGAGGACGCTTGCCA 1437  
QY 1435 CGGTGAGTTCATGACTGGGGTG 1457  
Db 1438 CGGTGGGGGTCAATGACTGGGGTG 1460

RESULT 21  
US-10-659-983A-4  
; Sequence 4, Application US/10659983A  
; Publication No. US2004015731A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovane, Timothy A  
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-284779  
; CURRENT APPLICATION NUMBER: US/10/659, 983A  
; CURRENT FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573,684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386,217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,219  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 1460  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: AOB Type C R5clone47 16D rDNA  
US-10-659-983A-4

Query Match 84.1%; Score 1225.4; DB 7; Length 1460;  
Best Local Similarity 92.1%; Pred. No. 0;  
Matches 1348; Conservative 0; Mismatches 106; Indels 9; Gaps 5;

QY 1 ATTTGAACGCTGGCGGCTGCTTTACATGCAAGTCAACGCGACGCGATGCTTGAT 60  
Db 1 ATTTGAACGCTGGCGGCTGCTTTACATGCAAGTCAACGCGAG--CGGGGCTTCGCG 58  
QY 61 CTGGTGGCGATGGCGGAGCGGTGATGATCGGACGTATCCAGAGAGGGGGTGA 120  
Db 59 CTGGCGGCGATGGCGGAGCGGTGATGATCGGACGTATCCAGAGAGGGGGATGA 118  
QY 121 ACGCATGAAAGATGTGCTAATACCGCATATACTTAAGAGGAAAGCAGGGGATCGAAA 180  
Db 119 ACGCATGAAAGATGTGCTAATACCGCATAT-CTCTAGAGAAAGCAGGGGATCGCAA 177  
QY 181 GACCTTGCGCTTTTGAAGCGCGGATGCTGATTAAGTAACTGGTGGGTTAAAGGCTTAC 240  
Db 178 GACCTTGCGCTTTAAGAGCGCGGATGCTGATTAAGTAACTGGTGGGTTAAAGGCTTAC 237  
QY 241 CAAGGCGAGATGATGATGTGCTGAGAGAGCAGCAGGACCACTGGGACTGAGACAG 300  
Db 238 CAAGGCGAGATGATGATGTGCTGAGAGAGCAGCAGCAGCAGCAGCTGGGACTGAGACAG 297  
QY 301 GCCCAGACTCTTAACGGAGGAGCAGAGTGGGAAATTTTGAACAATGGCGCAGACCTGATC 360  
Db 298 GCCCAGACTCTTAACGGAGGAGCAGAGTGGGAAATTTTGAACAATGGCGCAGACCTGATC 357  
QY 361 CAGCAATGCGCGTGAAGTGAAGAGGCTTCGGGTTTAAAGCTCTTTCACTGAGAGAA 420  
Db 358 CAGCAATGCGCGTGAAGTGAAGAGGCTTCGGGTTTAAAGCTCTTTTGTGAGAGAA 417  
QY 421 AAGGTTAAGTAAATATGATGACCATGACGATGACAGAGAAAGACGCGGCTAAC 480

Db 418 AAGAAATCATGATGAATTAATTATGATGACGTAAGTACAGAGAAAGACCGGCTAAC 477  
QY 481 TACGTGCGCAGCAGCGCGGTAAATACGTAGGGTGCAGAGCTTAAATGGAAATTCTGGGCGT 540  
Db 478 TACGTGCGCAGCAGCGCGGTAAATACGTAGGGTGCAGAGCTTAAATGGAAATTCTGGGCGT 537  
QY 541 AAAAGGTGCGCAGCGCGCTTGAAGTCAAGTGAATAATCCCGGGCTTAACTGGGAAT 600  
Db 538 AAAAGGTGCGCAGCGCGCTTGAAGTCAAGTGAATAATCCCGGGCTTAACTGGGAAT 597  
QY 601 TCGCTTTGAAACTACAAAGCTAGAGTGGGAGAGGAGGTGAATTCATGTGTGACAG 660  
Db 598 TCGCTTTGAAACTACAAAGCTAGAGTGGGAGAGGAGGTGAATTCATGTGTGACAG 657  
QY 661 TGAATGCTTAAAGATATGGAAGAACATGATGCGGAGGAGCGCTCCCTGGGTAACTACT 720  
Db 658 TGAATGCTTAAAGATATGGAAGAACATGATGCGGAGGAGCGCTCCCTGGGTGACACT 717  
QY 721 GACGCTCATGACAGAAAGCGGGAGAGCAAGAGATTAGATACCTGGTAGTCCAGCGC 780  
Db 718 GACGCTCATGACAGAAAGCGGTGGAGAGCAAGAGATTAGATACCTGGTAGTCCAGCGC 777  
QY 781 CTAAACGATGTCAACTAGTTGTGGGCTTATTA--GGCTTGATACGAACTAACCGCT 838  
Db 778 CTAAACGATGTCAACTAGTTGTGGGCTTATTAAGATTGTGTAACGTAGCTAACCGCT 837  
QY 839 GAAAGTTGACCGGCTGGGAGGTACGCTCCGCAATTAATACTAAAGAAATTGAACGGGAGC 898  
Db 838 GAAAGTTGACCGGCTGGGAGGTACGCTCCGCAATTAATACTAAAGAAATTGAACGGGAGC 897  
QY 899 CCGCAGACGGGTGATTAATGTGATTAATTCATGACAGCAAGCAAAACCTTAACCTACCC 958  
Db 898 CCGCAGACGGGTGATTAATGTGATTAATTCATGACAGCAAGCAAAACCTTAACCTACCC 957  
QY 959 TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTGC--TTGCGGAACGCTAACACAG 1015  
Db 958 TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTGC--TTGCGGAACGCTAACACAG 1017  
QY 1016 TGTGATGATGGCTGTGCTGACGCTGCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGG 1075  
Db 1018 TGTGATGATGGCTGTGCTGACGCTGCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGG 1077  
QY 1076 CAACCTGTGCTAATTAATTCGATCAAT--TGGTTGGGCACTTAATGAGACTCGGGTGC 1134  
Db 1078 CAACCTGTGCTAATTAATTCGATCAAT--TGGTTGGGCACTTAATGAGACTCGGGTGC 1137  
QY 1135 AAACCGAGAGAGTGGGATGACGTCAAGTCTCAATGACCTTAATGGGTAGGGCTTCA 1194  
Db 1138 AAACCGAGAGAGTGGGATGACGTCAAGTCTCAATGACCTTAATGGGTAGGGCTTCA 1197  
QY 1195 ACGTAATCAATGGCGGCTGACAGAGGTTGCCAACCCCGGAGGGGAGCTAATTCAGAA 1254  
Db 1198 ACGTAATCAATGGCGGCTGACAGAGGTTGCCAACCCCGGAGGGGAGCTAATTCAGAA 1257  
QY 1255 AGCGGTGATGCGGATCGAGTCTGCAATCGACTCGTGAAGTGGAAATCGGTAGT 1314  
Db 1258 AGCGGTGATGCGGATCGAGTCTGCAATCGACTCGTGAAGTGGAAATCGGTAGT 1317  
QY 1315 AATCGCGATCAAGATGCGCGGTGAATAGTTCCCGGGTCTTGTACACACGCGCGTCA 1374  
Db 1318 AATCGCGATCAAGATGCGCGGTGAATAGTTCCCGGGTCTTGTGTACACACGCGCGTCA 1377  
QY 1375 CACCATGGAGTGGGTTTCCAGAAAGCAGTAACTTAACCGTAAAGAGAGGGCTTGCCA 1434  
Db 1378 CACCATGGAGTGGGTTTCCAGAAAGCAGTAACTTAACCGTAAAGAGAGAGCTTGCCA 1437  
QY 1435 CGGTGAGTTCATGACTGGGGTG 1457  
Db 1438 CGGTGAGTTCATGACTGGGGTG 1460

RESULT 22  
US-10-168-337A-10



```
; Sequence 10, Application US/10168337A
; Publication No. US20030170654A1
; GENERAL INFORMATION:
; APPLICANT: Croceci, Gregory R.; Tyson, Gene W.; Hugenholtz, Philip; and Blackall,
; APPLICANT: Linda L.
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyporphosphate
; TITLE OF INVENTION: Accumulating Organisms in Wastewater
; FILE REFERENCE: 002367PC/KF
; CURRENT APPLICATION NUMBER: US/10/168,337A
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 10
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Propionibacter pelophilus
US-10-168-337A-10

Query Match      82.5%; Score 1201.6; DB 6; Length 1485;
Best Local Similarity 89.8%; Pred. No. 0;
Matches 1313; Conservative 0; Mismatches 144; Indels 5; Gaps 2;

Qy      1 ATTGAAGCTGCGCGCATGCTTTACATGCAAGTCGAAACGGACGACGATGCTTGAT 60
Db      8 ATTGAAGCTGCGCGCATGCTTTACATGCAAGTCGAAACGGACGATGCTTGAT 67
Qy      61 CTGTGCGGAGTGGCGGACGGGTGAGTATCATCGAAACGTATCCAGAAAGGGGGGTA 120
Db      68 CTGATGCGGAGTGGCGGACGGGTGAGTATCATCGAAACGTATCCAGAAAGGGGGGTA 127
Qy      121 ACGCATGAAAGATGCTCTAATACCGCATATCTCTAAGAGAAAGAGGGGATCGAAA 180
Db      128 ACGTAGCGAAAGTATCGCTAATACCGCATATCTCTGAGCAAGAAAGGGGGATCGCAA 187
Qy      181 GACCTTGGCTTTTGGAGCGCGCGCATGCTCTAGTTAGTATGAGTGGGGGTAAAGGCTAC 240
Db      188 GACCTTGGCTTTTGGAGCGCGCGCATGCTCTAGTTAGTATGAGTGGGGGTAAAGGCTAC 247
Qy      241 CAAAGCGACATCATAGTATGCTGTCGAGAGACGACCAACCTGCGGACTGAGACAG 300
Db      248 CAAAGCGACATCATAGTATGCTGTCGAGAGATGATCCGCCACACTGCGGACTGAGACAG 307
Qy      301 GCCCAGACTCTCTACCGGAGGACAGCATGCGGGAATTTTGAACATGCGCGCAAGCTGATC 360
Db      308 GCCCAGACTCTCTACCGGAGGACAGCATGCGGGAATTTTGAACATGCGCGCAAGCTGATC 367
Qy      361 CAGCAATCCCGCTGAGAGGAAAGGCTTGGGGTGTAAAGCTCTTCACTGCGAGAGA 420
Db      368 CAGCCTATCCCGCTGAGAGGAAAGGCTTGGGGTGTAAAGCTCTTCACTGCGAGAGA 427
Qy      421 AAAGGTATCGGTAATAATCGTGAACCATGACCGGTATGACAGAGAGACACCGGCTAAC 480
Db      428 AATGCAAGCTCTAATAGCTGTGATGACCGGTATGACAGAGAGACACCGGCTAAC 487
Qy      481 TACGTGCAAGACCGCGGTAAATACGTAGAGGTGCAAGCGTTAATGGAATTAATCTGGCGT 540
Db      488 TACGTGCAAGACCGCGGTAAATACGTAGAGGTGCAAGCGTTAATGGAATTAATCTGGCGT 547
Qy      541 AAAGGTGCGGAGCGGCTTGTAGTACAGATGTAATCCCGGGCTTAATCTGGGAT 600
Db      548 AAAGGTGCGGAGCGGCTTGTAGTACAGATGTAATCCCGGGCTTAATCTGGGAT 607
Qy      601 TGCGTTTGAACCTAACAAAGCTAGAGTGGCAGAGGAGTGAATTCATGCTGTAGAG 660
Db      608 GGCCTTTGAGACTGACAGCGCTAGAGTGAAGAGGAGGAGTGAATTCACGTGTAGAG 667
Qy      661 TGAATATGCTAGAGATATGAAAGAAACATGATGCGAAGCGACGCTCTGCGTTAACT 720
Db      668 TGAATATGCTAGAGATATGAAAGAAACATGATGCGAAGCGACGCTCTGCGTTAACT 727
Qy      721 GACGCTCATGCAAGAAAGGTGGGAGCAAAAGATTAAGATACCTGTGTAGTCCAGCGC 780
Db      728 GACGCTCATGCAAGAAAGGTGGGAGCAAAAGATTAAGATACCTGTGTAGTCCAGCGC 787
```

```
Qy      781 CTAAGCATGTCACACTAGTTGTTG--GCCTTATAGGCTTGGTAACGAAGCTAACGCGT 838
Db      788 CTAAGCATGTCACACTAGTTGTTGAGGAGGTAAACCTTATAGTGTGACTAGCGCT 847
Qy      839 GAAGTTGACCGCTGCGGAGTACGCTCGCAAGATTAATACTCAAGAGATTTGACGGGAC 898
Db      848 GAAGTTGACCGCTGCGGAGTACGCGCGCAAGGCTAAATCTCAAGAGATTTGACGGGAC 907
Qy      899 CCGCAGACGGGTGAGATTAATGAGATTAATTCAGATGCAAGCGGAAAACTTACTTACC 958
Db      908 CCGCAGACGGGTGAGATTAATGAGATTAATTCAGATGCAAGCGGAAAACTTACTTACC 967
Qy      959 TTGACATGTCAGAAATTTTCTAGAGATAGATTGTC---TTGCGGAACGCTAACAGAG 1015
Db      968 TTGACATGTCAGAAATCTTGAAGATTAATGAGATGCGGAAAGGAGCTGAAACAGAG 1027
Qy      1016 TGCTGCAATGCTGTCGACGCTGTCGTGAGATGTTGGGTTAAGTCCGCAACGAGCG 1075
Db      1028 TGCTGCAATGCTGTCGTCACGCTGTCGTGAGATGTTGGGTTAAGTCCGCAACGAGCG 1087
Qy      1076 CAACCTTGTCTAATTAATTCGCATCATTTGTTGGGCACTTTAATGAGATGCGGTCACA 1135
Db      1088 CAACCTTGTCTAATTAATTCGCATCATTAATGTTGGGCACTTTAATGAGATGCGGTCACA 1147
Qy      1136 AACCGAGAGAGGTGGGATGACGTCAAGTCCATGCGCCCTAATGGGTAGGGCTTACA 1195
Db      1148 AACCGAGAGAGGTGGGATGACGTCAAGTCCATGCGCCCTAATGGGTAGGGCTTACA 1207
Qy      1196 CGTAATCAATATGCGCGCTACAGAGGGTTGCAACCCCGAGGGGAGCTAATCTCAGAAA 1255
Db      1208 CGTAATCAATATGCGCGTTTCAAGAGGTTGCAACCCCGAGGGGAGCTAATCTCAGAAA 1267
Qy      1256 GCGCGTGTAGTCCGATCGGATCGGATCTGCAACTGCACTCCGTAAGTGGAAATGCTTGTGTA 1315
Db      1268 GCGCGTGTAGTCCGATCGGATCGGATCTGCAACTGCACTGCAATGAGTGGAAATGCTTGTGTA 1327
Qy      1316 ATCGCGATACAGCATGTCGCGGTAAATACGTTCCCGGGCTTGTAAACACCGCGGTAC 1375
Db      1328 ATCGCGATACAGCATGTCGCGGTAAATACGTTCCCGGGCTTGTAAACACCGCGGTAC 1387
Qy      1376 ACCATGGAGAGTGGTTTACCAAGAGAGGATGCTAATCCGTAAGAGAGGCGCTTGCAC 1435
Db      1388 ACCATGGAGAGGAGTTTTCGCCAAGAGTATGAGCTTAACCGCAAGAGAGGCGCTTACAC 1447
Qy      1436 GGTGAGATTCAATGACTGGGGTG 1457
Db      1448 GCGGGGTTGCTGACTGGGGTG 1469

RESULT 23
US-10-168-337A-6
; Sequence 6, Application US/10168337A
; Publication No. US20030170654A1
; GENERAL INFORMATION:
; APPLICANT: Croceci, Gregory R.; Tyson, Gene W.; Hugenholtz, Philip; and Blackall,
; APPLICANT: Linda L.
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyporphosphate
; TITLE OF INVENTION: Accumulating Organisms in Wastewater
; FILE REFERENCE: 002367PC/KF
; CURRENT APPLICATION NUMBER: US/10/168,337A
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 6
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:
; OTHER INFORMATION: Polyporphosphate-accumulating organism
US-10-168-337A-6
```

Query Match 81.7%; Score 1190.8; DB 6; Length 1460;  
Best Local Similarity 89.8%; Pred. No. 0;  
Matches 1313; Conservative 0; Mismatches 142; Indels 7; Gaps 3;

QY 1 ATTGAAGCTGGCGGCGGCTTTTACATGCAAGTCGCAACGCGACGACGCGATGCTTGCAAT 60  
Db 1 ATTAAACGCTGGCGGCGGCTTTTACATGCAAGTCGCAACGCGACGCGGCGCAAC-- 58  
QY 61 CTGTGGCGGATGGCGGACGGGTGAGTAATGATCGGAACGTATCCGAAGAGGGGGTA 120  
Db 59 CTGTGGCGGATGGCGGACGGGTGAGTAATGATCGGAACGTATCCGAAGAGGGGGTA 118  
QY 121 ACCGATCGAAAGATGTCCTAATACCGATATACCTTAAGAGGAAACAGGGGATCGAAA 180  
Db 119 ACCGAGGAAAGCTACCGTAAATACCGATATCTTGAAGCAAGAAACAGGGGATCGCAA 178  
QY 181 GACCTTGCGCTTTTGGAGCGGCGCGATGCTGATTTAGCTAGTTGGTGGGCTTAAAGCCCTAC 240  
Db 179 GACCTTGCGCTTTTGGAGCGGCGCGATGCTGATTTAGCTAGTTGGTGGGCTTAAAGCCCTAC 238  
QY 241 CAAGCGACGATCATGATGTTGCTGAGAGGACGACGACCACTGGGACTGAGACAG 300  
Db 239 CAAGCGACGATCATGATGTTGCTGAGAGGATGATCCGACACTGGGACTGAGACAG 298  
QY 301 GCCCAGACTCTCTACGGGAGGACGATGGGGAATTTTGAACAATGGCGGAGGCTGATC 360  
Db 299 GCCCAGACTCTCTACGGGAGGACGATGGGGAATTTTGAACAATGGCGGAGGCTGATC 358  
QY 361 CAGCAATGCCGCTGATGAGAAAGCCCTTCGGGTTGTAAGCTCTTTCAGTCGAGAGA 420  
Db 359 CAGCAATGCCGCTGATGAGAAAGCCCTTCGGGTTGTAAGCTCTTTCAGTCGAGAGA 418  
QY 421 AAAGTTACCGTAAATATATCTGACCCATGACCGTATGACAGAAAGACACCGGCTTAC 480  
Db 419 AAATGCTTGGGTTAATATCTCTGATGATGATACCGTATGACAGAAAGACACCGGCTTAC 478  
QY 481 TAGGTGCGACAGCGCGGATATATCGTAAAGGTCGACGCTTAATCGGAATTTACTGGGCGT 540  
Db 479 TAGGTGCGACAGCGCGGATATATCGTAAAGGTCGACGCTTAATCGGAATTTACTGGGCGT 538  
QY 541 AAAGGTCGCGACGCGGCTTGTAAAGTCAGATGGAATCCCGGCTTAACTGCGGAAT 600  
Db 539 AAAGGTCGCGACGCGGCTTGTGTAAAGTCAGATGGAATCCCGGCTTAACTGCGGAAT 598  
QY 601 TGGCTTTGAAACTCAAAAGCTAGTGTGCGAGGAGGAGGTGGAATTCATGTGTACAG 660  
Db 599 TGCATTTGAACTGCAAGACTGAGGAGTGTGCGAGGAGGAGGTGGAATTCACGTGTACAG 658  
QY 661 TGAATGCGTAAAGATATGAAAGCATGATGCGAAGGACGCTCTGGGTTAACTACT 720  
Db 659 TGAATGCGTAAAGATATGAAAGCATGATGCGAAGGACGCTCTGGGTTAACTACT 718  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAATACCTGTGATGTCAGGCC 780  
Db 719 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAATACCTGTGATGTCAGGCC 778  
QY 781 CTAAACGATGTCAACTAGTTGTTGG--GCCTTATTAAGCTTTGTAAGCAAGCTTACGCGT 838  
Db 779 CTAAACGATGTCAACTAGTTGTTGGAGGTTAAACCTTTAAGTCCGTAGCTAACGCGT 838  
QY 839 GAAGTTGACCGGCTGGGAGATTAAGGTGCGAAGATTAATCAAAAGAAATTTGACGCGGAC 898  
Db 839 GAAGTTGACCGGCTGGGAGATTAAGGTGCGAAGGCTTAACAAAGAAATTTGACGCGGAC 898  
QY 899 CCGCACAAGCGGTGATATGATGATTAATTCGATGCAACGCAAAAACCTTACTTACCC 958  
Db 899 CCGCACAAGCGGTGATGATGATTAATTCGATGCAACGCAAAAACCTTACTTACCC 958  
QY 959 TTGACATGTAAGCAATTTCTTAAGATTAATGAGTCT---TGGGAAACGCTTACACAGG 1015  
Db 959 TTGACATGTAAGCAATTTCTTAAGATTAATGAGTCTCGCAAGAGAACTTGAACACAGG 1018  
QY 1016 TGCTGATGCTGTCTGCTGCTGTCTGTCTGTGAGATGTTGGGTTAAAGTCCGCAACGACG 1075

|||||  
Db 1019 TGCTGATGCTGTCTGCTGCTGTCTGTGAGATGTTAGTCCCGCAACGACGCG 1078  
QY 1076 CAACCTTTGATTAATTTATCCCATCATTTTGGTGGGCACTTTAATGAGATGCGCGTGACA 1135  
Db 1079 CAACCTTTGATTAATTTATCCCATCATTTTGGTGGGCACTTTAATGAGATGCGCGTGACA 1138  
QY 1136 AACCGGAGGAAGGTGGGATGACGTCAAGTCCCTCACTGCGCCCTTATGGTGAAGGCTTACCA 1195  
Db 1139 AACCGGAGGAAGGTGGGATGACGTCAAGTCCCTCACTGCGCCCTTATGGTGAAGGCTTACCA 1198  
QY 1196 CGTAAATACATGACGCGTACAGAGGTTGCAACCCGCGAGGGGAGAGCTAATCTCAGAAA 1255  
Db 1199 CGTAAATACATGATGCTGATCAGAGGTTGCAACCCGCGAGGGGAGGCAATCTCAGAAA 1258  
QY 1256 GCGCGTCTGATGTCGGATGCGAGTCTGCAACTCTGATCCGTGAAGTCGGAATCTGTAAT 1315  
Db 1259 GCGGATCTGATGTCGGATGCGAGTCTGCAACTCTGATCCGTGAAGTCGGAATCTGTAAT 1318  
QY 1316 ATCGCGATCAGCATGTCGGGTTGAATACGTTCCCGGCTTGTACACGCGCGCTCAC 1375  
Db 1319 ATCGCGATCAGCATGTCGGGTTGAATACGTTCCCGGCTTGTACACGCGCGCTCAC 1378  
QY 1376 ACCATGGAGTGGGTTTCCAGAAAGCAGGTAGTCTAACCTTAAGAGGCGCTTGCAC 1435  
Db 1379 ACCATGGAGGCGGTTCTGCGCAAGATGTTAGCTTAACCGCAAGAGGCGATTACAC 1438  
QY 1436 GGTGAGATTCAATGACTGGGGTG 1457  
Db 1439 GGCAGGGTTCTGTGACTGGGGTG 1460

RESULT 24  
US-10-168-337A-8  
; Sequence 8, Application US//10168337A  
; Publication No. US20030170654A1  
; GENERAL INFORMATION:  
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholz, Philip; and Blackall,  
; APPLICANT: Linda L.  
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyphosphate  
; FILE REFERENCE: 002367PC/KP  
; CURRENT APPLICATION NUMBER: US/10/168,337A  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 1459  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:  
US-10-168-337A-8

Query Match 80.8%; Score 1177.8; DB 6; Length 1459;  
Best Local Similarity 89.8%; Pred. No. 0;  
Matches 1300; Conservative 1; Mismatches 143; Indels 7; Gaps 3;

QY 12 GCGGATGCTTTTACATGATGCAAGTCGAAACGCGACGATGCTTGCATCTGTGGCGAG 71  
Db 11 GCGGATGCTTTTACATGATGCAAGTCGAAACGCGACGATGCTTGCATCTGTGGCGAG 68  
QY 72 TGGCGGACGGGTGAGTAATGATCGGAACGTATCGAAGAGGAGGAGTAAAGCATCGAAA 131  
Db 69 TGGCGGACGGGTGAGTAATGATCGGAACGTATCGAAGAGGAGGAGTAAAGCATCGAAA 128  
QY 132 GATGCTTAATACCGGATTAATCTTAAGAGGAAAGCAGGGGATGGAAGACCTTGCGCT 191  
Db 129 GTTAGCGTAATACCGGATTAATCTTAAGAGGAAAGCAGGGGATGGAAGACCTTGCGCT 188  
QY 192 TTTGGAGCGCGGATGCTGATTAAGTGTGTGGGTTAAAGGCTTACCAAGCGACGA 251

Db 189 CTGGAGCGCGCGATGTCGAGTTAGCTAGTTGGGTAAAGGCTTACCAAGGCCAACA 248  
QY 252 TCAGTAGTTGGTCTGAGAGACGACGACCACTGAGACTGAGACACGGGCCCAAGCTCC 311  
Db 249 TCCGTAGCGGGTCTGAGAGATGATCCGCCACACTGGAGCTGAGACCGGCCCAAGCTCC 308  
QY 312 TACGGAGGCGAGAGTGGGGAATTTTGGACATGGCGGACGCTGATCCAGCATGCCG 371  
Db 309 TACGGAGGCGAGAGTGGGGAATTTTGGACATGGCGGACGCTGATCCAGCATGCCG 368  
QY 372 CGTAGTAGAAGAGCGCTTGGGGTGTGAAGCTCTTCACTGAGAGAAAGGTTACGG 431  
Db 369 CGTAGTAGAAGAGCGCTTGGGGTGTGAAGCTCTTCACTGAGAGAAAGGTTACGG 428  
QY 432 TAATATATCGTGACCCATCGGATATGACAGAAAGACCGCGCTAACTACGTGCGACG 491  
Db 429 TTAATATCCTGTGTAGATGACGGTATCCCGAATTAAGAAAGACCGCGCTAACTACGTGCGACG 488  
QY 492 AGCCGCGGTAAATAGTAGAGGTGCAAGCGTTAATCGGAATTAATCGGCGTAAAGGGTCCG 551  
Db 489 AGCCGCGGTAAATAGTAGAGGTGCGAGCGTTAATCGGAATTAATCGGCGTAAAGGGTCCG 548  
QY 552 AGCGCGCTTGTAGTCAAGTGTGAATCCCGGGCTTAACTGGGAATTCGCTTGAA 611  
Db 549 AGCGCGTGTGTAGTCAAGTGTGAATCCCGGGCTTAACTGGGAATTCGCTTGAA 608  
QY 612 CTACAAAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTAGAGCATGAAATGCGTA 671  
Db 609 CTGCAAGGCTGAGATTTTGGCAGAGGGGGGTGGAATTCAGTGTAGAGCATGAAATGCGTA 668  
QY 672 GAGATATGAAAGACATCATGATGCGGAAGGACGCGCTTGGGTAACTGACGCTCATGC 731  
Db 669 GAGATATGAGGAACAACGATGCGGAAGGACGCGCTTGGGTAACTGACGCTCATGC 728  
QY 732 ACGAAGCGTGGGAGCAACAGGATTAAGTACCTGTGTAGTCCAGCCCTTAAACGAT 791  
Db 729 ACGAAGCGTGGGAGCAACAGGATTAAGTACCTGTGTAGTCCAGCCCTTAAACGAT 788  
QY 792 CAACTAGTTGTGG--GCTTATTAAGGCTGTGTAAAGGATTAAGGCTGAGTGAACG 849  
Db 789 CAACTAGTTGTGGAGGCTTAAACCTTTAGTGTGTGTAAAGGCTGAGTGAACG 848  
QY 850 CTTGGGAGTACGCTGCGAAGATTAACTCAAGAAATTTGACGCGGACCGCAAGCG 909  
Db 849 CTTGGGAGTACGCGCGGAGCTTAACTCAAGAAATTTGACGCGGACCGCAAGCG 908  
QY 910 GTGATTAATGTGATTAATTCGATGCAACGGAAGAACTTAACTTCCCTTGAACATGAG 969  
Db 909 GTGATTAATGTGATTAATTCGATGCAACGGAAGAACTTAACTTCCCTTGAACATGCA 968  
QY 970 CGAATTTTCTAGAGATGATTAAGTCT--TCGGGAGCGTTAAACAGAGTCTGATGCG 1026  
Db 969 GGAATTCCTGAAGAGATTTGGAGTGTCTGCAAGAGAGCTTAAACAGAGTCTGATGCG 1028  
QY 1027 TGTGTCAGCTCGTGTCTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAACCTTGTCT 1086  
Db 1029 TGTGTCAGCTCGTGTCTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAACCTTGTCT 1088  
QY 1087 ATTAATTCATCAATTTGGTGGGCACTTTAATGAGACTGCGCGGTGACAAACCGAGGAA 1146  
Db 1089 ATTAATTCATCAATTTAGTTGGGCACTTTAATGAGACTGCGCAAGTGAACAAACCGAGGAA 1148  
QY 1147 GGTGGGATGACGTCAAGTCTCTCAATGGGCTTAAATGGGATTTCAACAGGTAATCAAT 1206  
Db 1149 GGTGGGATGACGTCAAGTCTCTCAATGGGCTTAAATGGGATTTCAACAGGTAATCAAT 1208  
QY 1207 GCGCGGTACAGAGGGTTGCAACCGCGAGGGGAGCTAATCTCAAGAAAGCGCGTGTAG 1266  
Db 1209 GGTGCGTCCAGAGGGTTGCAACCGCGAGGGGAGGCAATCTCAAGAAAGCGGATGTAG 1268  
QY 1267 TCCGAGTGGAGTCTTGCACTTGAAGTTCGGAATGCTTAAATGCTTAAATGCTGAGATCA 1326  
Db 1269 TCCGAGTGGAGTCTTGCACTTGAAGTTCGGAATGCTTAAATGCTTAAATGCTGAGATCA 1328

QY 1327 GGATGTCGGGGAATTAAGTTCGGGGTCTTGTACACACCGCGCTCAACCATGGAGT 1386  
Db 1329 GGAATGCGGGTAATACCTTCCGGGTCTTGTACACACCGCGCTCAACCATGGAGT 1388  
QY 1387 GGGTTTACAGAGAGGATGATCTTAACGTAAGAGGGCGCTTGCACGCTGAGATTTCA 1446  
Db 1389 GGGTTTACAGAGAGGATGATCTTAACGTAAGAGGGCGCTTGCACGCGGAGGTTTCG 1448  
QY 1447 TGAATGGGGTG 1457  
Db 1449 TGAATGGGGTG 1459  
  
RESULT 25  
US-10-168-337A-4  
; Sequence 4, Application US/10168337A  
; Publication No. US20030170654A1  
; GENERAL INFORMATION:  
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholz, Philip; and Blackall,  
; APPLICANT: Linda L.  
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyphosphate  
; FILE REFERENCE: 002367PC/KF  
; CURRENT APPLICATION NUMBER: US/10/168,337A  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1460  
; TYPE: DNA  
; ORGANISM: Rhodocyclops tennis  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1045, 1315  
; OTHER INFORMATION: n = unknown  
US-10-168-337A-4  
  
Query Match 80.7%; Score 1176; DB 6; Length 1460;  
Best Local Similarity 89.1%; Pred. No. 0;  
Matches 1303; Conservative 0; Mismatches 152; Indels 7; Gaps 3;  
  
QY 1 ATGAAGCGTGGGGGATGCTTAACTGATGCAAGTGAACGCGGACGAGCTTGCAT 60  
Db 1 ATGAAGCGTGGGGGATGCTTAACTGATGCAAGTGAACGCGGACGAGCTTGCAT-- 58  
QY 61 CTGGTGGCGAGTGGCGGACGGGTGATGATGATCGGAAGTTCAGAAAGAGGGGGTA 120  
Db 59 CTGGTGGCGAGTGGCGGACGGGTGATGATGATCGGAAGTTCAGAAAGTGGGGGTA 118  
QY 121 ACGCATGAAAGATGTGCTAATACCGCATATCTTAAGAGGAAAGCAGGGATTCGAA 180  
Db 119 ACGTAAAGAAAGTATACCTAATACCGCATATCTGTAGAGGAAAGCAGGGATTCGAA 178  
QY 181 GACCTTGGCTTTTGGAGCGCGCGATGCTGATTAAGTGAATGGGTAAAGGCTTAC 240  
Db 179 GACCTTGGCTTTTGGAGCGCGCGATGCTGATTAAGTGAATGGGTAAAGGCTTAC 238  
QY 241 CAAGCGAGATGATGATGTTGTCTGAGAGGACGACGAGCACTGGGACTGAGACG 300  
Db 239 CAAGCGAGATGATGATGTTGTCTGAGAGGACGACGAGCACTGGGACTGAGACG 298  
QY 301 GCCAGACTCTTAACGAGAGGACGAGTGGGAAATTTTGGACATGGGCGCAAGCTTATC 360  
Db 299 GCCAGACTCTTAACGAGAGGACGAGTGGGAAATTTTGGACATGGGCGCAAGCTTATC 358  
QY 361 CAGCAATCCGCGTGAATGAAGAGGCTTCGGGTGTAAGCTCTTCACTGAGAGAA 420  
Db 359 CAGCAATCCGCGTGAATGAAGAGGCTTCGGGTGTAAGCTCTTCACTGAGAGAA 418  
QY 421 AAAGTTACGGTAAATATGCTGACCGATGACGATATGACAGAGAAAGCACCGGCTTAC 480  
Db 419 AAATGCTCAGATTAATACCTGATGATGAGGTTACCGAAGAAAGCACCGGCTTAC 478





Db 959 TTGACATGTCAGGAAATCTTGAAGATTTAGGAGTCCCGAAAGGAACTGAACACAG 1018  
 Qy 1016 TGTGTCATGCTGTCTGTCAGCTGTGTCTGAGATGTTGGTTAACTCCCGCAACGACG 1075  
 Db 1019 TGTGTCATGCTGTCTGTCAGCTGTGTCTGAGATGTTGGTTAACTCCCGCAACGACG 1078  
 Qy 1076 CAACCTTGTCTAATTTGTCATCTTTGGTGGGCACTTATATGAGCTCCGGTGACA 1135  
 Db 1079 CAACCTTGTCTAATTTGTCATCTTTGGTGGGCACTTATATGAGCTCCGGTGACA 1138  
 Qy 1136 AACCGGAGGAGGTGGGGATGAGTCAAGTCTCATGAGCCCTTATGGGTAGGCTTACAC 1195  
 Db 1139 AACCGGAGGAGGTGGGGATGAGTCAAGTCTCATGAGCCCTTATGGGTAGGCTTACAC 1198  
 Qy 1196 CGTAATACATGCGCGTCACAGAGGTTGCGCAACCCGCGAGGGGAGCTAATCTCAGAAA 1255  
 Db 1199 CGTCATACATGCTGCTCGTCAGAGGTTGCGCAACCCGCGAGGGGAGCTAATCTCAGAAA 1258  
 Qy 1256 GCGCGTCGTAGTCCGATCGGAGTCTGCAACTGCACTCCGTGAAGTCGGAATGCTAAGTA 1315  
 Db 1259 GCGGATCGTATGTCGCGATGTCAGTCTGCAACTGCACTGCAATGGAATCGGAATGCTAAGTA 1318  
 Qy 1316 ATGCGCGATCAGATGTCGCGGATGCAAGTCCCGGCTTGTATACACACCGCCGTCAC 1375  
 Db 1319 ATGCGCGATCAGATGTCGCGGATGCAAGTCCCGGCTTGTATACACACCGCCGTCAC 1378  
 Qy 1376 ACCATGGAAGTGGGTTTCACAGAGGAGTGTCTAACCGTAAGAGGGCGTGGCCAC 1435  
 Db 1379 ACCATGGAAGTGGGTTTCACAGAGGAGTGTCTAACCGTAAGAGGGCGTGGCCAC 1438  
 Qy 1436 GGTGAATTCATGACTGGGGTG 1457  
 Db 1439 GGCAGGATTCGTGACTGGGGTG 1460

RESULT 28  
 US-10-029-397A-38  
 / Sequence 38, Application US/10029397A  
 / Publication No. US20030175709A1  
 / GENERAL INFORMATION:  
 / APPLICANT: MURPHY, GEORGE L.  
 / APPLICANT: WHITLEY, J. PENN  
 / TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS  
 / FILE REFERENCE: AMB1:0760S  
 / CURRENT APPLICATION NUMBER: US/10/029.397A  
 / NUMBER OF SEQ. ID NOS: 73  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ. ID NO 38  
 / LENGTH: 1532  
 / TYPE: DNA  
 / ORGANISM: Bordetella bronchiseptica  
 US-10-029-397A-38

Query Match 78.9%; Score 1149.2; DB 6; Length 1532;  
 Best Local Similarity 88.2%; Pred. No. 1,1e-309;  
 Matches 1287; Conservative 0; Mismatches 163; Indels 10; Gaps 3;

Qy 1 ATGGAACGCTGGGGGACGCTTTACATGCAAGTCAAGCGGACGACGAGATGCTTGCAT 60  
 Db 29 ATGGAACGCTGGGGGACGCTTTACATGCAAGTCAAGCGGACGACGAG--GCTTGGC 86  
 Qy 61 CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGAACGTATCCAGAAAGGGGGGTA 120  
 Db 87 CTGGTGGCGAGTGGCGGAAACGGGTGAGTAATGTAATCGAAACGTGCCAGTAGCGGGGAGTA 146  
 Qy 121 ACCGATGGAAGATGTCTAATACCGATATATCTTAAAGAGGAAACAGAGGATCGAAA 180  
 Db 147 ACTACGGAAGAGGTGCTAATACCGATATACCGCTTACGGGGGAAACGGGGGAGCTTGC 206  
 Qy 181 GACCTTGGCTTTTGAAGCGGCGAGTGTGATTAAGCTAGTGGTGGGTAAAGGCTTAC 240

Db 207 GGCCTGCACTATTGAGAGCGCCGATATTCGATTAGTGTGTGGGTAAACGGCTTAC 266  
 Qy 241 CAAGCGACGATACATTAAGTTGGTCTGAGAGGACGACCAAGCCACATCTGGGACTGAGACAG 300  
 Db 267 CAAGCGACGATACATTAAGTTGGTCTGAGAGGACGACCAAGCCACATCTGGGACTGAGACAG 326  
 Qy 301 GCCCAAGCTCTTACGGGAGGAGCAGAGTGGGGAATTTGGACAAATGGGCGCAAGCTGATC 360  
 Db 327 GCCCAAGCTCTTACGGGAGGAGCAGAGTGGGGAATTTTGGACAAATGGGCGCAAGCTGATC 386  
 Qy 361 CAGCAATGCGCGTGAATGAGAGAGGCTTGGGTTGTAAGCTTTTCACTGAGGAGAGA 420  
 Db 387 CAGCAATGCGCGTGAATGAGAGAGGCTTGGGTTGTAAGCTTTTCACTGAGGAGAGA 446  
 Qy 421 AAGGTACGATTAATATTCGTGACCCATGAGGATGAGCAGAGAGGACCGGCTAAC 480  
 Db 447 AACGCGACGGGTAAATATTCGTGACATGACGAGTACCTGCAAGATTAAGCACCGGCTAAC 506  
 Qy 481 TACGTGCGACAGCGCGGTAAATACGTAGGGTGCAGAGCTTAAATCGGAATTAATCGGCGT 540  
 Db 507 TACGTGCGACAGCGCGGTAAATACGTAGGGTGCAGAGCTTAAATCGGAATTAATCGGCGT 566  
 Qy 541 AAGGGTGGCAGGCGGCTTGTAAATGTCAGATGTGAATCCCGGCTTAACTGGGAAT 600  
 Db 567 AAGGGTGGCAGGCGGCTTGTAAATGTCAGATGTGAATCCCGGCTTAACTGGGAAT 626  
 Qy 601 TGCCTTTGAATCTAACAAGCTAGTGTGAGAGGAGGTGGAATTCATGTGTAGCAG 660  
 Db 627 TGCATTTTAACTACCGGCTAGAGTGTGAGAGGAGGTGGAATTCGCTGTAGCAG 686  
 Qy 661 TGAATGCTAGAGATATGAGAGAAATCATGATGTCGAGAGGAGGCTCTGGGTTAACTACT 720  
 Db 687 TGAATGCTAGAGATATGAGAGAAATCATGATGTCGAGAGGAGGCTCTGGGTTAACTACT 746  
 Qy 721 GACGTCATGACAGAAAGCTGGGGAGCAAAACAGATTGATATCCCTGTATCCAGCGC 780  
 Db 747 GACGTCATGACAGAAAGCTGGGGAGCAAAACAGATTGATATCCCTGTATCCAGCGC 806  
 Qy 781 CTAAACGATGTAACATTAATGTTGGGCTTAAATTAAGCTTGTATGCAAGCTAAACGCGTGA 840  
 Db 807 CTAAACGATGTAACATTAATGTTGGGCTTAAATTAAGCTTGTATGCAAGCTAAACGCGTGA 866  
 Qy 841 AGTTGACCGCTGGGGAGTACGCTGTCAGAGATTAATACTCAAGAAATGACGGGGACCC 900  
 Db 867 AGTTGACCGCTGGGGAGTACGCTGTCAGAGATTAATACTCAAGAAATGACGGGGACCC 926  
 Qy 901 GCAACAAGCGGTGATTAATGATGATTAATTCGATGCAACCGGAAAACTTAACTTAACCTT 960  
 Db 927 GCAACAAGCGGTGATTAATGATGATTAATTCGATGCAACCGGAAAACTTAACTTAACCTT 986  
 Qy 961 GACATGTAGCAATTTTCTAGAGATTAATTAAGTACT--TGGGAAACGCTTAACACAGGTG 1017  
 Db 987 GACATGTCTGGAATCCCGAAGAAATTTGGAGTGTCTCGAAGAGAACCGGAACAGAGTG 1046  
 Qy 1018 CTGCATGCTGTCTGTCAGCTGCTGTCGTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1077  
 Db 1047 CTGCATGCTGTCTGTCAGCTGCTGTCGTGAGATGTTGGTTAAGTCCGCAACGAGCGCA 1106  
 Qy 1078 ACCCTTGTCAATTAATTCGATCAATTTGGTTGGGCACTTTAATGAGACTGCGGTGACAAA 1137  
 Db 1107 ACCCTTGTCAATTAATTCGATCAATTTGGTTGGGCACTTTAATGAGACTGCGGTGACAAA 1161  
 Qy 1138 CCGGAGAAAGTGGGGATGACAGTCAAGTCTCAATGCGCTTAAATGGGTAGGGCTTTCACAG 1197  
 Db 1162 CCGGAGAAAGTGGGGATGACAGTCAAGTCTCAATGCGCTTAAATGGGTAGGGCTTTCACAG 1221  
 Qy 1198 TAATCAATGCGCTGTCACAGAGGTTGCCAACCCGCGAGGGGAGCTAAATCTCAGAAAGC 1257  
 Db 1222 TCATTAATGTCGGGACAGAGGCTGCCAACCCGCGAGGGGAGCTAAATCTCAGAAAGC 1281  
 Qy 1258 GCGTGTAAATGTCGGATGCGAGTCTGCAACTCGATCCGTGAAGTCCGAATCGTATTAAT 1317  
 Db 1282 CGATCGTATGTCGGATGCGAGTCTGCAACTCGATCCGTGAAGTCCGAATCGTATTAAT 1341





```

; TITLE OF INVENTION: VARIATION DETECTION AND DISCOVERY
; FILE REFERENCE: SEQ-2073-UT
; CURRENT APPLICATION NUMBER: US/10/723,365B
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,895
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 32
; LENGTH: 1496
; TYPE: DNA
; ORGANISM: Bordetella petrii
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (821)
; OTHER INFORMATION: a, c, g, or t
US-10-723-365B-32

Query Match      78.3%; Score 1140.6; DB 9; Length 1496;
Best Local Similarity 88.0%; Freq. No. 2,7e-307;
Matches 1279; Conservative 0; Mismatches 165; Indels 10; Gaps 3;

QY      7 CGCTGCGCGCATCTTTACATGATGCAAGTCGAACGGCAGCAGCATGCTTGACATCTGTG 66
DB      1 CGCTAGCGGGATGCTTTACATGATGCAAGTCGAACGGCAGCAGCGA--CTTCGGTCTGGG 58
QY      67 GCGAGTGGCGGACGGGTGAGTAATGCAATCGGAACGTAATCCAGAAAGGGGGGTAACGCAT 126
DB      59 GCGAGTGGCGGAACGGGTGAGTAATGTAATCGGAACGTAATCCAGAAAGGGGGGTAATCAAG 118
QY      127 CGAAATATGCTTAATACCGCATATATCTTAAGAGAAAGAAAGGGGATCGAAAGCCTT 186
DB      119 CGAAATCTTAAGTAATACCGCATATACCGCTTAAGAGAAAGGGGGGAAAGCTTCGGGCTTC 178
QY      187 GCGCTTTTGAAGCGCGCATGCTGTAAGTAAGTCTGAGTGGTGAAGGCTTACCAAGGC 246
DB      179 GCACTTTTGAAGCGCGCATGCTGTAAGTAAGTCTGAGTGGTGAAGGCTTACCAAGGC 238
QY      247 GACGATCAGTAAGTGTCTGAGAGAGACGACACGACCACTGGGACTGAGACACGCGCCAG 306
DB      239 GACGATCCGTAAGTGTGTTGAGAGAGACGACACGACCACTGGGACTGAGACACGCGCCAG 298
QY      307 ACTCTTACCGGAGGACGACGATGCGGGAATTTTGAACATGAGGCGCAAGCCTTACCGAGCA 366
DB      299 ACTCTTACCGGAGGACGACGATGCGGGAATTTTGAACATGAGGCGCAAGCCTTACCGAGCA 358
QY      367 TGGCGGCTGAGTAAGAGAGCCTTCGAGGTTTGAAGCTCTTTCAGTCGAGAAAGAAAGGT 426
DB      359 TCCCGGCTGAGTAAGAGAGCCTTCGAGGTTTGAAGCACTTTTGGCAGAAAGAAAGAGC 418
QY      427 TACGTAATTAATCGTACCATGACCGTATCGACAGAAAGACACCGGCTTAATCTACGTG 486
DB      419 TCTGGCTAATACCTGGGGCAATGACGCTACCTGCAAGAAATTAACACCGGCTTAATCTACGTG 478
QY      487 CCAAGCAGCGCGGTAATACGTAAGGTGCAAGCCTTAATCGGAATTTCTGGGCGCTAAAGG 546
DB      479 CCAAGCAGCGCGGTAATACGTAAGGTGCAAGCCTTAATCGGAATTTCTGGGCGCTAAAGG 538
QY      547 TGGCAGCGCGGCTTGTAAAGTCAGATGTGAATCCCGGCGCTTAACCTTGGAATTCGCT 606
DB      539 TGGCAGCGCGGCTTGGAAAGAAAGATGTGAATCCCAAGGCTTAACTTTGAATTCGCT 598
QY      607 TGAATACCAAAAGCTAAGAGTGTGCAAGAGGAGGTGAATTCATGTGTAGCAGTGAAT 666
DB      599 TTTAACTACCGGCTAAGAGTGTGCAAGAGGAGGTGAATTCGCGGTGTAGCAGTGAAT 658
QY      667 GCGTAAGATATGGAAGACATGATGGCGAAGCAGCCTCTCGGGTAAACATGACGCT 726
DB      659 GCGTAAGATATGCGGAGGAACACCGATGGCGAAGCAGCCTCTCGGGTAAACATGACGCT 718
QY      727 CATGACGAAGAGCGTGGGAGCAAAACAGATTAGATACCTGTGTAAGTCCAGCGCCCTTAAC 786
DB      719 CATGACGAAGAGCGTGGGAGCAAAACAGATTAGATACCTGTGTAAGTCCAGCGCCCTTAAC 778
```

```

QY      787 GATGCACTAGTGTGGGCGCTTATAGGCTTGTGAACGAAGCTTAACCGGTAAGTTGA 846
DB      779 GATGCACTAGTGTGGGCGCTTATAGGCTTGTGAACGAAGCTTAAACCGGTAAGTTGA 838
QY      847 CCGCTGGGAGTAAGGTGCAAGATTAATACTCAAGAAATTGAAGGGGACCCGACAA 906
DB      839 CCGCTGGGAGTAAGGTGCAAGATTAATACTCAAGAAATTGAAGGGGACCCGACAA 898
QY      907 GCGGTGATTAATGTAATTAATTCATGACCAACGCAAAAACCTTACCTTACCATG 966
DB      899 GCGGTGATTAATGTAATTAATTCATGACCAACGCAAAAACCTTACCTTACCATG 958
QY      967 TAGCAATTTTCTAGATAGATTAGTGT---TGGGAACGCTTAACACAGGTGCGAT 1023
DB      959 TCTGAATGCGCAAGAGATTGCGAGTGTGCGAAGAGAACCGAACACAGGTGCGAT 1018
QY      1024 GCGTGTGCAAGCTGCTGCTGAGATGTTGGGTTAAGTCCGCAACGACGCAACCTT 1083
DB      1019 GCGTGTGCAAGCTGCTGCTGAGATGTTGGGTTAAGTCCGCAACGACGCAACCTT 1078
QY      1084 GTCAATTAATGCAATTAATTTGTTGGGCACTTTAATGACACTGCGGTGACAAACCGGAG 1143
DB      1079 GTCAATTAATGCAATTAATTTGTTGGGCACTTTAATGACACTGCGGTGACAAACCGGAG 1133
QY      1144 GAAAGTGGGGAATGAGTCAAGTCAATGAGCCCTTAAGGTTAGGGCTTACACGTAATAC 1203
DB      1134 GAAAGTGGGGAATGAGTCAAGTCAATGAGCCCTTAAGGTTAGGGCTTACACGTAATAC 1193
QY      1204 AATGCGCGCTACAGAGGTTGCAACCCGCAAGGGGAGTAATCTCAAGAAACCGGCTG 1263
DB      1194 AATGCGCGGGAATGAGTCAAGTCAATGAGCCCTTAAGGTTAGGGCTTACACGTAATAC 1253
QY      1264 TAGTCGGATGCGAGTCTGCAACTGACTCCGTGAAGTCCGAATGCTTAATGCGGGA 1323
DB      1254 TAGTCGGATGCGAGTCTGCAACTGACTCCGTGAAGTCCGAATGCTTAATGCGGGA 1313
QY      1324 TCAGCATGTCGGGTAAGTAAGTCCCGGCTTGTACACACCGGCGGCAACCATGGG 1383
DB      1314 TCAGCATGTCGGGTAAGTAAGTCCCGGCTTGTACACACCGGCGGCAACCATGGG 1373
QY      1384 AGTGGGTTTACACAGAGCAGTATCTTAACCGTAAGAGGCGGCTTGCACGTAAGAT 1443
DB      1374 AGTGGGTTTACACAGAGTATCTTAACCGTAAGAGGCGGCTTGCACGTAAGAT 1433
QY      1444 TCATGACTGGGCTG 1457
DB      1434 TCATGACTGGGCTG 1447

RESULT 31
US-10-875-161-1
; Sequence 1, Application US/10875161
; Publication No. US2005009151A1
; GENERAL INFORMATION:
; APPLICANT: Chase, Matthew
; APPLICANT: Claydon, Robert
; APPLICANT: Landis, Bryan
; APPLICANT: Banerjee, Amit
; TITLE OF INVENTION: Method for the Stereoselective Synthesis and Enantiomeric
; FILE REFERENCE: SO-3262-2-PR-US
; CURRENT APPLICATION NUMBER: US/10/875,161
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1530
; TYPE: DNA
; ORGANISM: Varioloxax paradoxus
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(1530)
```



OTHER INFORMATION: 16S rRNA gene (rDNA) from Varilovorax paradoxus  
FEATURE:  
NAME/KEY: misc\_difference  
LOCATION: (1) .. (1530)  
OTHER INFORMATION: alignment 0.75% different from 16S rRNA gene (rDNA) of Varilovorax  
OTHER INFORMATION: paradoxus  
US-10-875-161-1

Query Match 78.0%; Score 1136; DB 8; Length 1530;  
Best Local Similarity 87.5%; Pred. No. 5,2e-306;  
Matches 1277; Conservative 1; Mismatches 176; Indels 6; Gaps 3;

1 ATTGAACGCTGGCGGATCTTACACATGCAAGTGAACGAGACGGAATCTTGAT 60  
24 ATTGAACGCTGGCGGATCTTACACATGCAAGTGAACGAGACGGAATCTTGAT 81  
61 CTGGTGGCAGATGGCGGAGCGGATGATGATCGGAACGTAACGAAAGAGGGGAT 120  
82 CTGGCGGCGAGTGGCGAAGCGGATGATGATCGGAACGTAACGTAACGTAACG 141  
121 ACGCATGAAAGATGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGATGAAA 180  
142 ACGCAGGAAAGCTGTGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGATG 201  
181 GACCTTGGCTTTGGAGGGCGGATGCTGATTAAGTATGTTGGTGGGTAAGGCT 240  
202 GACCTTGGCGGATGAGGCGCGGATGAGGATTAAGTATGTTGGTGGGTAAGG 261  
241 CAAGGCGAGATGATGATGCTGAGAGAGAGACGACGACACTGGGACTGAGACAG 300  
262 CAAGCTTGGATTTGATGCTGAGAGAGAGACGACGACACTGGGACTGAGACAG 321  
301 GCCCAGACTCTTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
322 GCCCAGACTCTTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 381  
361 CAGCAATGCGCGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
382 CAGCCATGCGCGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 441  
421 AAGGTTAAGGTAATTAATGATGACCGATGACCGATGACCGATGACCGATGAC 480  
442 AAGGTTAAGGTAATTAATGATGACCGATGACCGATGACCGATGACCGATGAC 501  
481 TACGTCCAGCAGCAGCGGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
502 TACGTCCAGCAGCAGCGGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 561  
541 AAGGTTGCGCAGCGGCTTTGATGATGATGATGATGATGATGATGATGATGAT 600  
562 AAGGTTGCGCAGCGGCTTTGATGATGATGATGATGATGATGATGATGATGAT 621  
601 TGGCTTGAACCTAAGATGATGATGATGATGATGATGATGATGATGATGATG 660  
622 TGGCTTGAACCTAAGATGATGATGATGATGATGATGATGATGATGATGATG 681  
661 TGAATGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
682 TGAATGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 741  
721 GACGCTCATCAGAAAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
742 GACGCTCATCAGAAAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 801  
781 CTAAACGATGCACTGTTGTTGGGCTTAATGATGATGATGATGATGATGATG 840  
802 CTAAACGATGCACTGTTGTTGGGCTTAATGATGATGATGATGATGATGATG 861  
841 AGTTGACCGCTGGGAGAGTACGATCGCAAGATTAACTCAAGAGATTGACGGG 900  
862 AGTTGACCGCTGGGAGAGTACGATCGCAAGATTAACTCAAGAGATTGACGGG 921  
901 GCAACAGCGGTGATTAATGATGATTAATGATGATTAATGATGATTAATGAT 960

922 GCACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 981  
961 GACATGACGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGAT 1017  
982 GACATGACGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGAT 1041  
1018 CTGATGACGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGAT 1077  
1042 CTGATGACGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGAT 1101  
1102 ACCCTTGCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1160  
1138 CCGAGAGAGAGTGGGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
1161 CCGAGAGAGAGTGGGATGATGATGATGATGATGATGATGATGATGATGAT 1220  
1198 TAATCAATGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1257  
1221 TCATCAATGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1280  
1258 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317  
1281 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1340  
1318 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377  
1341 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1400  
1378 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1437  
1401 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1460  
1438 TGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1497  
1461 CAGGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1530

RESULT 32  
US-10-029-397A-41  
Sequence 41, Application US/10029397A  
Publication No. US20030175709A1  
GENERAL INFORMATION:  
APPLICANT: MURPHY, GEORGE L.  
APPLICANT: WHITLEY, J. PENN.  
TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING RNA POPULATIONS  
FILE REFERENCE: AMB:076US  
CURRENT APPLICATION NUMBER: US/10/029,397A  
CURRENT FILING DATE: 2002-03-19  
NUMBER OF SEQ ID NOS: 73  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 41  
LENGTH: 1535  
TYPE: DNA  
ORGANISM: Burkholderia cepacia  
US-10-029-397A-41

Query Match 77.9%; Score 1135.2; DB 6; Length 1535;  
Best Local Similarity 87.5%; Pred. No. 8,6e-306;  
Matches 1278; Conservative 0; Mismatches 173; Indels 9; Gaps 3;

1 ATTGAACGCTGGCGGATCTTACACATGCAAGTGAACGAGACGGAATCTTGAT 60  
29 ATTGAACGCTGGCGGATCTTACACATGCAAGTGAACGAGACGGAATCTTGAT 88  
61 CTGGTGGCAGATGGCGGAGCGGATGATGATCGGAACGTAACGAAAGAGGGGAT 120  
89 CTGGTGGCAGATGGCGGAGCGGATGATGATCGGAACGTAACGAAAGAGGGGAT 148  
121 ACGCATGAAAGATGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATGAAA 180

149 GCCCGCGAAGCCGGATTAAATACCGCATACGATCGAAGTAAAGCGGGGACCTTCG 208  
QY 181 GACCTTGCGCTTTTGGAGCGCCGATGCTGATTAAGCTAGTGGTGGGTAAAGCCCTAC 240  
Db 209 GGCCTCGCGCTAATAGG-GTTTGGCGATGAGCTGATTAAGTGGTGGGTAAAGCCCTAC 267  
QY 241 CAAGCGCAGATCACTAGTGGTCTGAGAGCAGCAGCCCACTGGGACTGAGACAG 300  
Db 268 CAAGCGCAGATCACTAGTGGTCTGAGAGCAGCAGCCCACTGGGACTGAGACAG 327  
QY 301 GCCCAGACTCTTACCGGAGGCGAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCTTC 360  
Db 328 GCCCAGACTCTTACCGGAGGCGAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCTTC 387  
QY 361 CAGCAATGCCCGGTGAGTGAAGAAGCCTTCGGGTTTGAAGCTCTTTCAGTCGAGAA 420  
Db 388 CAGCAATGCCCGGTGAGTGAAGAAGCCTTCGGGTTTGAAGCTCTTTCAGTCGAGAA 447  
QY 421 AAAGCTTACGGTAATTAATCGTACCCCATGACGGTATGACAGAGAACCGGCTTAC 480  
Db 448 AATCCCTGGCTTAATACATCGCGGAGATGACGGTATGACCGAGAAATAGCACCGGCTTAC 507  
QY 481 TACGTGCAGCAGCCCGGCTAATACGTAGGGTGCAGCGCTTAATCGGAATTACTGGGCGT 540  
Db 508 TACGTGCAGCAGCCCGGCTAATACGTAGGGTGCAGCGCTTAATCGGAATTACTGGGCGT 567  
QY 541 AAAGGTGCGCAGCGCGCTTGTAACTCAGATGTGAATCCCCGGGCTTAACCTGGGAAT 600  
Db 568 AAAGGTGCGCAGCGCGCTTGTAACTCAGATGTGAATCCCCGGGCTTAACCTGGGAAT 627  
QY 601 TGCGTTTAACTACAAAGCTAGTGTGGGAGGAGGAGGAGTTCATGTTGAGCAG 660  
Db 628 TGCAATGTGTAATGAGCTAGAGTATGACAGAGGGGGTGAATTCACGTATGACAG 687  
QY 661 TGAAATGCGTATGAGATATGAGAAACATCGATGCGAAGCGACCTCTCGGCTTAACCT 720  
Db 688 TGAAATGCGTATGAGATATGAGAAACATCGATGCGAAGCGACCTCTCGGCTTAACCT 747  
QY 721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAAGTACCTGTGATGCAAGCC 780  
Db 748 GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAAGTACCTGTGATGCAAGCC 807  
QY 781 CTAAACGATGCACTAGTGTGGGCTTATTAAGCTGTGATGCAAGCTTACCTGTA 840  
Db 808 CTAAACGATGCACTAGTGTGGGCTTATTAAGCTGTGATGCAAGCTTACCTGTA 867  
QY 841 AGTTGACCGCTGGGAGTACCGGTGCAAGATTAACCTCAAGGAATTTGACGGGAGCC 900  
Db 868 AGTTGACCGCTGGGAGTACCGGTGCAAGATTAACCTCAAGGAATTTGACGGGAGCC 927  
QY 901 GCACAAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTTACCTT 960  
Db 928 GCACAAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTTACCTT 987  
QY 961 GACATGTAAGCAATTTCTAGAGATTAATTAAGTGTCT---TGCGGAAAGCTAACACAGGTG 1017  
Db 988 GACATGTAAGCAATTTCTAGAGATTAATTAAGTGTCTCGAAGAGAAACCGGCGCAAGGTG 1047  
QY 1018 CTGCATGCTGCTGCTCAGCTCGTGTCTGAGATGTTGGGTTAAAGTCCGCAACGACGCA 1077  
Db 1048 CTGCATGCTGCTGCTCAGCTCGTGTCTGAGATGTTGGGTTAAAGTCCGCAACGACGCA 1107  
QY 1078 ACCCTTGCTAATTAATTTGCCATCATTTGGTGGGCACTTTAATGACATGCGCGGTGACAA 1137  
Db 1108 ACCCTTGCTAATTAATTTGGCTAAC-----GCAAGAGCACTCTAAGAGACATGCGCGGTGACAA 1162  
QY 1138 CCGAGGAAAGGTGGGAGTGAAGTCAAGTCTCAATGGCCCTTAATGGTGGGCTTACACG 1197  
Db 1163 CCGAGGAAAGGTGGGAGTGAAGTCAAGTCTCAATGGCCCTTAATGGGTTAAGGCTTACACG 1222  
QY 1198 TAAATCAATGCGCGCTTACAGAGGTTGCAACCCCGAGGCGGAGCTAATCTCAGAAAGC 1257  
Db 1223 TCATCAATGCTGCGAACAAGAGGTTGGCAACCCCGAGGCGGAGCTAATCTCAGAAAGC 1282

QY 1258 GCGTGTAGTCCGGATCGAGTCTGCAACTGCACTCCGTGAAGTCCGGAATCGTAGTAAT 1317  
Db 1283 CCAATGTAGTCCGATTCGACTCTGCAACTGCACTGAGTGAAGTCCGGAATCGTAGTAAT 1342  
QY 1318 CCGGATTCAGCATGTCGGGTGAATTAAGTTCGGGGTCTTTGTAACAACCGCCGTCAAC 1377  
Db 1343 CCGGATTCAGCATGTCGGGTGAATTAAGTTCGGGGTCTTTGTAACAACCGCCGTCAAC 1402  
QY 1378 CATGAGTGGGTTTCAACCAAGCAGTACTTAACCGTAAGGAGGCGCTTCCACCG 1437  
Db 1403 CATGAGTGGGTTTCAACCAAGCAGTACTTAACCGCAAGAGAGCGGTCAACACG 1462  
QY 1438 TGAGATTCATGACTGGGGTG 1457  
Db 1463 TAGATTCATGACTGGGGTG 1482

RESULT 33  
US-10-029-397A-39  
; Sequence 39, Application US/10029397A  
; Publication No. US20030175709A1  
; GENERAL INFORMATION:  
; APPLICANT: MURPHY, GEORGE L.  
; APPLICANT: WHITLEY, J. BENN  
; TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS  
; FILE REFERENCE: AMBI-076US  
; CURRENT APPLICATION NUMBER: US/10/029,397A  
; CURRENT FILING DATE: 2002-03-19  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 1485  
; TYPE: DNA  
; ORGANISM: Bordetella parapertussis  
US-10-029-397A-39

Query Match 77.8%; Score 1134; DB 6; Length 1485;  
Best Local Similarity 87.9%; Pred. No. 1.9e-305;  
Matches 1285; Conservative 0; Mismatches 165; Indels 12; Gaps 4;

QY 1 ATTTGACGCTGGCGGATGCTTTACATGCAAGTGCAGCGGACGAGCTTGCAT 60  
Db 1 ATTTGACGCTGGCGGATGCTTTACATGCAAGTGCAGCGGACGAGCTTGCAT 58  
QY 61 CTGTGTGCGAATGCGCGAGCGGTGAGTAAATCGGAATCCGAAAGAGGGGGTA 120  
Db 59 CTGTGTGCGAATGCGCGAGCGGTGAGTAAATCGGAATCCGGAATCGGGGATA 118  
QY 121 ACGCATCGAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAA 180  
Db 119 ACTACGGAAGCGGTGCTAATACCGCATATATCGCTTACGGGGGAAAGCGGGGACTTTCCG 178  
QY 181 GACCTTGCGCTTTTGAAGCGGCGCATGTCTGATTAAGTGTGGTGGGTAAAGGCTTAC 240  
Db 179 GACCTTGCGCATATTGAGAGCGGCGCATTAAGCTTAAGTGTGGTGGGTAAAGGCTTAC 238  
QY 241 CAAGCGCAGATCACTAGTGTCTGAGAGAGCAACCAAGCTGAGGACTGAGACAG 300  
Db 239 CAAGCGCAGATCACTAGTGTCTGAGAGAGCAACCAAGCTGAGGACTGAGACAG 298  
QY 301 GCCCAGACTCTTACCGGAGGCGAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATC 360  
Db 299 GCCCAGACTCTTACCGGAGGCGAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATC 358  
QY 361 CAGCAATGCCCGGTGAGTGAAGAAGCCTTCGGGTTTGAAGCTCTTTCAGTCGAGAA 420  
Db 359 CAGCAATGCCCGGTGAGTGAAGAAGCCTTCGGGTTTGAAGCTCTTTCAGTCGAGAA 418  
QY 421 AAAGGTACGGTAATTAATGCTGACCCATGACCGTATGACAGAGAAAGCAGCGGCTTAC 480  
Db 419 AAAGGTACGGCTAATTAATCTGTGCAACTGACCGGTATCTGAGATTAAGCAGCGGCTTAC 478

QY 481 TACGTCCAGACCCCGGTAATAGTAGGGTCAAGCCTTAATCGGAATTAAGTGGCGT 540  
DB 479 TACGTCCAGACCCCGGTAATAGTAGGGTCAAGCCTTAATCGGAATTAAGTGGCGT 538  
QY 541 AAAGGTGGCGAGCGCGCTTGAAGTCAAGTGAATCCCGGGCTTAACCTGGGAAT 600  
DB 539 AAAGGTGGCGAGCGCGCTTGAAGTCAAGTGAATCCCGGGCTTAACCTGGGAAT 598  
QY 601 TGGCTTGAACCTCAACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTACAG 660  
DB 599 TGAATTTTAACTACCGGGCTAGAGTGTGTCAAGGGAGGTGGAATTCCTCGGTGTACAG 658  
QY 661 TGAATGCGTAGAGATATGGAAGAACATCGATGCGAAGGACGCTTCTGGGTAACTACT 720  
DB 659 TGAATGCGTAGAGATATGCGAGGAACACCGATGCGAAGGACGCTTCTGGGTAACTACT 718  
QY 721 GACGCTCATCAGAAAGCTGTGGGAGCAACAGAAATTAATCCCGGTAGTCCAGGCC 780  
DB 719 GACGCTCATCAGAAAGCTGTGGGAGCAACAGAAATTAATCCCGGTAGTCCAGGCC 778  
QY 781 CTAAAGATGCTCACTAGTTGTTGGGCTTATTAGGCTTGTGTAAAGAACTAAAGCGCTGA 840  
DB 779 CTAAAGATGCTCACTAGTTGTTGGGCTTATTAGGCTTGTGTAAAGAACTAAAGCGCTGA 838  
QY 841 AGTTGACCGCTGGGAGTACGGTCCGAGATTAACTCAAGAAATTGACGGGAGCC 900  
DB 839 AGTTGACCGCTGGGAGTACGGTCCGAGATTAACTCAAGAAATTGACGGGAGCC 898  
QY 901 GCAACAGCGGTGATTAATGTGATTAATTTGATGCAACGGGAAACCTTAACCTTACCCT 960  
DB 899 GCAACAGCGGTGATTAATGTGATTAATTTGATGCAACGGGAAACCTTAACCTTACCCT 958  
QY 961 GACATGAGGAATTTCTAGAGTATAGTGTCT---TCGGAGAGCTTAACAGAGTGT 1017  
DB 959 GACATGAGGAATTTCTAGAGTATAGTGTCT---TCGGAGAGCTTAACAGAGTGT 1018  
QY 1018 CTGATGAGGTGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
DB 1019 CTGATGAGGTGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078  
QY 1078 ACCCTTGTCAATTAATGCTCAATTTGGTGGGCACTTTAATGAGTCTGCGGTGAGCA 1137  
DB 1079 ACCCTTGTCAATTAATGCTCAATTTGGTGGGCACTTTAATGAGTCTGCGGTGAGCA 1133  
QY 1138 CCGGAGAGGTGGGAGTGAAGTCAAGTCTCATGAGCCTTAATGGGTAGGCTTACACAG 1197  
DB 1134 CCGGAGAGGTGGGAGTGAAGTCAAGTCTCATGAGCCTTAATGGGTAGGCTTACACAG 1193  
QY 1198 TAAATCAATGCGGTACAGAGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAAGC 1257  
DB 1194 TCAATCAATGCTCGGAGACAGAGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAAGC 1253  
QY 1258 GCGTGTAGTCCGAGTCCGAGTCTGCACTCTGCACTCTGTAAGTCCGTAATGCTAAT 1317  
DB 1254 GCGTGTAGTCCGAGTCCGAGTCTGCACTCTGCACTCTGTAAGTCCGTAATGCTAAT 1313  
QY 1318 CCGGATACAGATGTGCGGTGGAATAGTTCCTCGGTCTTGTAAACAACCGCGCTTACAC 1377  
DB 1314 CCGGATACAGATGTGCGGTGGAATAGTTCCTCGGTCTTGTAAACAACCGCGCTTACAC 1373  
QY 1378 CATGGAGTGGGTTTCAACAGAGGTAAGTCTAAACGCTAA--GGAAGGCGCTTGCAC 1435  
DB 1374 CATGGAGTGGGTTTCAACAGAGGTAAGTCTAAACGCTAA--GGAAGGCGCTTGCAC 1433  
QY 1436 GGTGAGATTCAATGACTGGGGTG 1457  
DB 1434 GGTGAGATTCAATGACTGGGGTG 1455

RESULT 34  
US-10-168-337A-3  
; Sequence 3, Application US/10168337A  
; Publication No. US20030170654A1

GENERAL INFORMATION:  
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholz, Philip; and Blackall,  
; APPLICANT: Linda L.  
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyphosphate  
; TITLE OF INVENTION: Accumulating Organisms in Wastewater  
; FILE REFERENCE: 002367PC/KP  
; CURRENT APPLICATION NUMBER: US/10/168,337A  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1478  
; TYPE: DNA  
; ORGANISM: Rhodocyclus purpureus  
; NAME/KEY: misc.feature  
; LOCATION: 78, 541, 761, 926, 1033, 1058, 1257  
; OTHER INFORMATION: n = unknown  
US-10-168-337A-3

Query Match 77.6%; Score 1130.6; DB 6; Length 1478;  
Best Local Similarity 87.4%; Pred. No. 1.6e-304;  
Matches 1269; Conservative 0; Mismatches 176; Indels 7; Gaps 3;

QY 1 ATTGAAGCTGGGCGATGCTTTACATGCAAGTGCAGACGCGACGCGATCTTGCAAT 60  
DB 29 ATTGAAGCTGGGCGATGCTTTACATGCAAGTGCAGACGCGATCTTGCAAT 86  
QY 61 CTGGTGGCGATGGCGGAGCGGTGAGTAATGATCGGAACGTAATCCAGAAAGGGGGTGA 120  
DB 87 CGCGCAACGATGTGCGCAACGCGGTGAGTAATGATCGGAACGTAATCCAGAAAGGGGGTGA 146  
QY 121 ACCGATCGAAAGTGTGCTAATACCGCATTAATCTTAAGAGAAAGCAGGAGATCGAAA 180  
DB 147 ACCGATCGAAAGTGTGCTAATACCGCATTAATCTTAAGAGAAAGCAGGAGATCGAAA 206  
QY 181 GACCTTGGCTTTTGAAGGGCGGATGCTGATTAAGTGTGATGTTGATGGGTAAAGCCTTAC 240  
DB 207 GCGCTTGGCTTTTGAAGGGCGGATGCTGATTAAGTGTGATGTTGATGGGTAAAGCCTTAC 266  
QY 241 CAAGCGACGATGATGATGCTGCTGAGAGAGCAACGACACCTGGGACTGAGACAGC 300  
DB 267 CAAGCGACGATGATGCTGCTGAGAGAGCAACGACACCTGGGACTGAGACAGC 326  
QY 301 GCCGAGCTCTTAACGAGGAGCAGCATGTTGGGAAATTTTGGCAATGGGCGCAAGCTTATC 360  
DB 327 GCCGAGCTCTTAACGAGGAGCAGCATGTTGGGAAATTTTGGCAATGGGCGCAAGCTTATC 386  
QY 361 CAGCATGCGCGGTGATGAAAGGCTTCCGGGTGTAAGCTTTCAGTGCAGAAAGA 420  
DB 387 CAGCATGCGCGGTGATGAAAGGCTTCCGGGTGTAAGCTTTCAGTGCAGAAAGA 446  
QY 421 AAAGGTAAAGGTAAATATGCTGACCCATGACCGGTATCGAAGAAAGACCGGCTTAC 480  
DB 447 AAAGGTAAAGGTAAATATGCTGACCCATGACCGGTATCGAAGAAAGACCGGCTTAC 506  
QY 481 TACGTGCGACGCGCGGTAAATAGTAGGGTGCAGACGTTAATCGAATTAATCTGGCGCT 540  
DB 507 TACGTGCGACGCGCGGTAAATAGTAGGGTGCAGACGTTAATCGAATTAATCTGGCGCT 566  
QY 541 AAAGGTGCGAGCGCGCTTGTAAAGTCAAGTGTGAATCCCGGGCTTAACTGGGAAT 600  
DB 567 AAAGGTGCGAGCGCGCTTGTGTAAAGTCAAGTGTGAATCCCGGGCTTAACTGGGAAT 626  
QY 601 TGGCTTTGAACCTCAACAAAGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660  
DB 627 TGGCTTTGTGACTCAACAGCTAGAGTACCGGAGAGGGGGGTGGAATTCATGTGTAGCAG 686  
QY 661 TGAATGCGTAGAGATATGGAAGAACATCGATGCGAAGGACGCTTCTGGGTAACTACT 720  
DB 687 TGAATGCGTAGAGATATGGAAGAACATCGATGCGAAGGACGCTTCTGGGTAACTACT 746  
QY 721 GACGCTCATCAGAAAGCTGTGGGAGCAACAGAAATTAATCCCGGTAGTCCAGGCC 780



```
QY 1018 CTGCATGGCTGTGCTGCACTGCTGCTGATGATGTTAGTTAGTCCCGACGAGCGCA 1077
DB 1045 CTGCATGGCTGTGCTGCACTGCTGCTGATGATGTTAGTTAGTCCCGACGAGCGCA 1104
QY 1078 ACCCTTGTCTTAATTGGCCATCAATTGGTTGGCACTTTAAATGAGCTGCGGTGACAA 1137
DB 1105 ACCCTTGTCTTAATTGGCTC-----TACGAAAGGCACTTAATGAGCTGCGGTGACAA 1159
QY 1138 CCGGAGGAAGGTGGGATGAGTCAAGTCCCTCAATGAGCTTATGGGTAGGGCTTCAACG 1197
DB 1160 CCGGAGGAAGGTGGGATGAGTCAAGTCCCTCAATGAGCTTATGGGTAGGGCTTCAACG 1219
QY 1198 TAAATACAAATGCGCGATACAGAGGGTTGCCAACCCGCGAGGGAGCTAATCTCAAGAA 1257
DB 1220 TCAATACAAATGCGCGATACAGAGGGTTGCCAACCCGCGAGGGAGCTAATCTCAAGAA 1279
QY 1258 GCGCTCGATGCTCGGATGAGAGTCTGCAATCTGCACTCCGTGAAGTCCGGAATCCGTA 1317
DB 1280 CCGTCTGATGCTCGGATGAGAGTCTGCAATCTGCACTCCGTGAAGTCCGGAATCCGTA 1339
QY 1318 CCGGATCAGATGCTCGGATGAGAGTCTGCAATCTGCACTCCGTGAAGTCCGGAATCCGTA 1377
DB 1340 CCGGATCAGATGCTCGGATGAGAGTCTGCAATCTGCACTCCGTGAAGTCCGGAATCCGTA 1399
QY 1378 CATGGAGTGGGTTTACCAAGACAGATGCTTAACCGTAAAGAGGGCGCTTGCCACGG 1437
DB 1400 CATGGAGTGGGTTTACCAAGACAGATGCTTAACCGTAAAGAGGGCGCTTGCCACGG 1459
QY 1438 TGAGA 1442
DB 1460 TAGGA 1464

RESULT 36
US-10-397-551-1
; Sequence 1, Application US/10397551
; Publication No. US20030190658A1
; GENERAL INFORMATION:
; APPLICANT: Beck, James J
; APPLICANT: Barnett, Charles J
; APPLICANT: Yarnall, Michele S
; APPLICANT: Zeltouni, Lilian
; TITLE OF INVENTION: Diagnostic for the detection of Acidovorax avenae subsp.
; TITLE OF INVENTION: Diagnostic for the detection of Acidovorax avenae subsp.
; FILE REFERENCE: 7000JUSNP
; CURRENT APPLICATION NUMBER: US/10/397,551
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/367,628
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Acidovorax avenae subsp. citrulli
US-10-397-551-1

Query Match 77.4%; Score 1127.6; DB 6; Length 1481;
Best Local Similarity 87.3%; Pred. No. 1.1e-303;
Matches 1274; Conservative 0; Mismatches 174; Indels 12; Gaps 3;
```

```
QY 181 GACCTTGCGCTTTTGGAGGCGCGATGCTGATTAAGCTAGTGGTGGGTTAAAGGCTTAC 240
DB 177 GCGCTTGGCGGAACGAGAGGCGCGATGCGAGTATAGTGGTGGGTTAAAGGCTTAC 236
QY 241 CAAGCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
DB 237 CAAGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 296
QY 301 GCCCAGACTTCTTACCGGAGGCGAGCACTGGGGAAATTTTGGCAATGGGCGCAACCTGATC 360
DB 297 GCCCAGACTTCTTACCGGAGGCGAGCACTGGGGAAATTTTGGCAATGGGCGCAACCTGATC 356
QY 361 CAGCAATGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
DB 357 CAGCAATGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 416
QY 421 AAAGTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 417 AAAGCTTCTTCTTAATTAAGGGGGGTCATGACGGTACCGTAAAGATTAAGCAGCGGCTAAC 476
QY 481 TACGTGCGAGCAGCGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 477 TACGTGCGAGCAGCGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 536
QY 541 AAAGGTCGCGCAGCGCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 537 AAAGCGTCGCGCAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 596
QY 601 TGGCTTGAATCTAACAAGCTTAAGTGTGTCAGAGGAGGATGGAATTCATGTGTAGCAG 660
DB 597 TGCATTTGTGATCTGATCTGCTGAGATGAGCAGAGGGGATGGAATTCGCGGTGTACAG 656
QY 661 TGAATGCGTGAATATGAGAGAAATCGATGAGCGAGCGAGCTCTCTGGGTTAACT 720
DB 657 TGAATGCGTGAATATGAGAGAAATCGATGAGCGAGCGAGCTCTCTGGGTTAACT 716
QY 721 GAGCTCATGACAGAAAGCTGGGAGGACAAACAGATTAATCCCGGTATGTCAGAGCC 780
DB 717 GAGCTCATGACAGAAAGCTGGGAGGACAAACAGATTAATCCCGGTATGTCAGAGCC 776
QY 781 CTAAGAGATGCACTGATGTTGGGCTTATTAAGCTTGTATGATGATGATGATGATGATGATG 840
DB 777 CTAAGAGATGCACTGATGTTGGGCTTATTAAGCTTGTATGATGATGATGATGATGATGATG 836
QY 841 AGTTGACCGGCTGGGAGTACGCTGCGCAAGATTAAATCTAAAGAAATGACGGGAGCC 900
DB 837 AGTTGACCGGCTGGGAGTACGCTGCGCAAGATTAAATCTAAAGAAATGACGGGAGCC 896
QY 901 GCACAGCGGTCGATTAATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 897 GCACAGCGGTCGATTAATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATG 956
QY 961 GACATGTAAGGAATTTCTAGAGATAGATTAGTCT--TCGGGACGCTAACACAGATG 1017
DB 957 GACATGTAAGGAATTTCTAGAGATAGATTAGTCT--TCGGGACGCTAACACAGATG 1016
QY 1018 CTGCATGGCTGTGCTGCACTGCTGCTGATGATGTTAGTTAGTCCCGACGAGCGCA 1077
DB 1017 CTGCATGGCTGTGCTGCACTGCTGCTGATGATGTTAGTTAGTCCCGACGAGCGCA 1076
QY 1078 ACCCTTGTCTTAATTGGCCATCAATTGGTTGGCACTTTAAATGAGCTGCGGTGACAA 1137
DB 1077 ACCCTTGTCTTAATTGGCTC-----GAGGACCTCTAATGGGACCTGCGGTGACAA 1131
QY 1138 CCGGAGGAAGGTGGGATGAGTCAAGTCCCTCAATGAGCTTATGGGTAGGGCTTCAACG 1197
DB 1132 CCGGAGGAAGGTGGGATGAGTCAAGTCCCTCAATGAGCTTATGGGTAGGGCTTCAACG 1191
QY 1198 TAAATACAAATGCGCGATACAGAGGGTTGCCAACCCGCGAGGGAGCTAATCTCAAGAA 1257
DB 1192 TCAATACAAATGCGCGATACAGAGGGTTGCCAACCCGCGAGGGAGCTAATCTCAAGAA 1251
```

[illegible]

DB	357	CAGCGATGCCCGCGTGCAAGATGAAAGCGCTTCGGGTGTAAACGCTTTTGACGGAACGA	416
QY	421	AAAGGTACCGGTAAATATATGTCGACCCATGACGGTATGCACAGAAGAACCGGCTTAAC	480
DB	417	AAAGCCTTCTTCTTAATAAAGGGGGGTGATGACGGTAAACCGTAAGATTAAGACACCGGCTAAC	476
QY	481	TACGTGCACAGCGCGGGGTAAATACGTAAAGGGTGCACGGTTAAATCCGGAATTAATCTGGGCGT	540
DB	477	TACGTGCACAGCGCGGGGTAAATCGTAAGGGTGCACGGTTAAATCGGAATTAATCTGGGCGT	536
QY	541	AAAGGTCGCGACGCGCGCTTGTAAAGTCAGATGTGAATATCCCGGGGCTTAACTGGGAAT	600
DB	537	AAAGCGTGCAGCGCGGGGTGATGTAAAGACAGATGTGAATATCCCGGGGCTTAACTGGGAAT	596
QY	601	TGCGTTTGAAACTAACAAAGCTGATGATGTGGCAGAGGGAGGTGGAAATTCATGTGTAGCAG	660
DB	597	TGCATTTGTGACTCATCTGCTGAGTGCACGCGCAGAGGGGGAATGTGAATTCGCGGTGTAGCAG	656
QY	661	TGAAATCCGTAAAGATATATGAAAGAACATTCGATGGCGAAGCGACGCTCGGGTTAAACCT	720
DB	657	TGAAATCCGTAAAGATATATGAAAGAACACCGATGGCGAAGCGAATCCCTGGGCTGTACT	716
QY	721	GACGCTATGACAGAAAGCGTGGGAGACAAACAGATTTAGATATACCTGTATGTCACGACC	780
DB	717	GACGCTATGACAGAAAGCGTGGGAGACAAACAGATTTAGATATACCTGTATGTCACGACC	776
QY	781	CTAAACGATGTCAACTAGTTGTTGGGCGCTTATTAAGCTTGTTAAACGAAGCTTAACGCTGA	840
DB	777	CTAAACGATGTCAACTAGTTGTTGGGCTTCACTGACTCACTGAAGCTTAACGAAGCTTAACGCTGA	836
QY	841	AGTTGACCGCTGGGGAGTACCGGTCCGAGATTTAAACTCAAGAGAAATTAACGGGAGACC	900
DB	837	AGTTGACCGCTGGGGAGTACGCGCGGACGGTTAAACTCAAGAGAAATTAACGGGAGACC	896
QY	901	GCACAAAGCGGAGTAAATATGTGATTTAATTCGATGCAACGCGAAGAAACCTTAACCTT	960
DB	897	GCACAAAGCGGAGTAAATATGTGATTTAATTCGATGCAACGCGAAGAAACCTTAACCTT	956
QY	961	GACATGTAGCGAATTTCTAGAGATTAAGTATGCT--TCGGGAACGCTTAACACAGGTG	1012
DB	957	GACATGTAGCGAATTTCTAGAGATTAAGTATGCTCTCGAAGAGAAACCGTAACACAGGTG	1016
QY	1018	CTGCATGGCTGTGCTGATCGTGTCCGAGATGTGGGTTAAAGTCCCGCAACGAGCGCA	1072
DB	1017	CTGCATGGCTGTGCTGATCGTGTGTGAGATGTGGGTTAAAGTCCCGCAACGAGCGCA	1076
QY	1078	AACCTTGTCATTAATGCGCATCTTTGGTGGGCACTTTAATGACACTGCCGTGACAA	1132
DB	1077	AACCTTGTCATTAATGTCGTAAGAA-----GGGCACTTAATTTGGGACTGCCGTGACAA	1136
QY	1138	CCGAGAGAGGTGGGAGTGAAGTCAAGTCTTCATAGGCCCTTAATGGGTAGGGCTTACACG	1192
DB	1132	CCGAGAGAGGTGGGAGTGAAGTCAAGTCTTCATAGGCCCTTAATGGGTAGGGCTTACACG	1196
QY	1198	TAAATTCATATGGCGGTACAGAGGGGTGGCAACCGCGCAGGGGGAGCTAAATCTCAGAAAGC	1252
DB	1192	TCAATTCATATGGCTGGTACAGAGGGGTGGCAACCGCGCAGGGGGAGCTAAATCTCAGAAAGC	1256
QY	1258	GCGTTCGTATGTCGCGATCGGAATCTGCAACTGCACTCCGTGAAAGTCGGAATCGTATTAAT	1312
DB	1252	CAGTTCGTATGTCGCGATCGGAATCTGCAACTGCACTCCGTGAAAGTCGGAATCGTATTAAT	1316
QY	1318	CGCGGATCAGATGTCGCGGTGAATATGCTTCCGGGCTTTGTACACACCGCCGTCACAC	1372
DB	1312	CGCGGATCAGATGTCGCGGTGAATATGCTTCCGGGCTTTGTACACACCGCCGTCACAC	1376
QY	1378	CATGGAGATGGGTTTCAACCAAGACAGTAACTTAAACGTAAAGAGGGCGCTTGACACGG	1432
DB	1372	CATGGAGATGGGTTTCTGCAGAAAGTAAAGTAACTTAAACGTAAAGAGGGCGCTTACACGG	1436
QY	1438	TGAGATTCACTAGCTGGGGTGG 1457	
DB	1432	CAGGGTTCTGTACTGGGGTGG 1451	

RESULT 38  
US-10-029-397A-42  
Sequence 42, Application US/10029397A  
Publication No. US20030175709A1  
GENERAL INFORMATION:  
APPLICANT: MURPHY, GEORGE L.  
APPLICANT: WHITLEY, J. PENN.  
TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS  
FILE REFERENCE: AMB1:076US  
CURRENT APPLICATION NUMBER: US/10/029,397A  
NUMBER OF SEQ ID NOS: 73  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 42  
LENGTH: 1488  
TYPE: DNA  
ORGANISM: Burkholderia mallei  
US-10-029-397A-42

Query Match 77.1%; Score 1123.6; DB 6; Length 1488;  
Best Local Similarity 87.1%; Pred. No. 1.5e-302;  
Matches 1271; Conservative 0; Mismatches 179; Indels 10; Gaps 3;  
Query 1 ATGAAGCTGCGCGCATGCTTTACACATGCAAGTCGACGCGACGCGATGCTTGAT 60  
3 ATTGAACGCTGGCGCATGCTTTACACATGCAAGTCGACGCGACGCGATGCTTGAT 60  
61 CTGGTGGCGAGTGGCGGACGCGTGAATGATGATCGGACGATTCGAGAGGGGGTGA 120  
61 CTGGTGGCGAGTGGTGAACGCGTGAATGATGATCGGACGATTCGAGTGGGGGATA 120  
121 ACCGATGGAAGATGCTCTAATCCGATATACCTTAAGGAGGAAAGCAGGGGATCGAA 180  
121 GCCCGGGAAGCGGATTAATACCGATACGATCTGAAGTGAAGCGGGGACCTTCG 180  
181 GACCTTGGCTTTTGAAGCGCGCATGCTGATTAAGTGGTGGGATTAAGGCGCTTAC 240  
181 GGCTTGGCTTTTGAAGCGCGCATGCTGATTAAGTGGTGGGATTAAGGCGCTTAC 240  
241 CAAGGCGACATGATGATGCTTGAAGGACGACGACGACCTGGAATGAGACAG 300  
241 CAAGGCGACATGATGATGCTTGAAGGACGACGACGACCTGGAATGAGACAG 300  
301 GCCGAGCTCTTAAGGAGGACGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
301 GCCGAGCTCTTAAGGAGGACGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
301 GCCGAGCTCTTAAGGAGGACGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
361 CAGCAATGCGCGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
361 CAGCAATGCGCGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
421 AAAGTTACGTTAATATGTAACCAATGCGGTATGACAGAGAGAGAGAGAGAGAGAG 480  
421 AATCATTTGCTAATATCCGAGTGAATGACGATGAGGAGGAGGAGGAGGAGGAGG 480  
481 TACGTGCGAGAGCGCGGTAAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
481 TACGTGCGAGAGCGCGGTAAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
541 AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
541 AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
601 TGCCTTTGAAATCAAAAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
601 TGCATGTTGACTGCGAGGAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
661 TGAATGCTGATGATGAG 720  
661 TGAATGCTGATGATGAG 720

Query 721 GACGCTATGACGAGAAAGCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
721 GACGCTATGACGAGAAAGCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
Query 781 CTAAACGATGCACTAATGTTGGGCTTTATTAAGCTTGGTGAAGAGCTAAGCGGTGA 840  
781 CTAAACGATGCACTAATGTTGGGCTTTATTAAGCTTGGTGAAGAGCTAAGCGGTGA 840  
Query 841 AGTTGACCCCTGGGAGAGTACGCTGCGCAAGATTAAGGAGGAGGAGGAGGAGGAGG 900  
841 AGTTGACCCCTGGGAGAGTACGCTGCGCAAGATTAAGGAGGAGGAGGAGGAGGAGG 900  
Query 901 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
901 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Query 961 GACATGAGCAATTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1017  
961 GACATGAGCAATTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1017  
Query 1018 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
1018 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
Query 1021 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1021 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Query 1078 ACCCTTGTCAATTAATGCGATCATATTTGTTGGGCACTTAATGAGACTGCGGTGACAA 1137  
1078 ACCCTTGTCAATTAATGCGATCATATTTGTTGGGCACTTAATGAGACTGCGGTGACAA 1137  
Query 1081 ACCCTTGTCTTAATGTTGCTAC----GCAAGACACTTAAGAGACTGCGGTGACAA 1135  
1081 ACCCTTGTCTTAATGTTGCTAC----GCAAGACACTTAAGAGACTGCGGTGACAA 1135  
Query 1138 CCGAGAGAGTGGGAGTACGTCAGTCCCTCATGCGCCCTTAATGAGGCTTCAACAG 1197  
1138 CCGAGAGAGTGGGAGTACGTCAGTCCCTCATGCGCCCTTAATGAGGCTTCAACAG 1197  
Query 1136 CCGAGAGAGTGGGAGTACGTCAGTCCCTCATGCGCCCTTAATGAGGCTTCAACAG 1195  
1136 CCGAGAGAGTGGGAGTACGTCAGTCCCTCATGCGCCCTTAATGAGGCTTCAACAG 1195  
Query 1198 TAATCAATGAGCGGTCAGAGGAGGTTCCCAACCCGAGAGGAGGAGGAGGAGG 1257  
1198 TAATCAATGAGCGGTCAGAGGAGGTTCCCAACCCGAGAGGAGGAGGAGGAGGAGG 1257  
Query 1196 TCATCAATGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1255  
1196 TCATCAATGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1255  
Query 1258 GCGTCGATGCGGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1317  
1258 GCGTCGATGCGGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1317  
Query 1256 CGATCGATGCGGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1315  
1256 CGATCGATGCGGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1315  
Query 1318 CCGGATCAGCATGCTGCGGTGAATGCTTCCGAGTCTTGTACACACCGCCGTCACAC 1377  
1318 CCGGATCAGCATGCTGCGGTGAATGCTTCCGAGTCTTGTACACACCGCCGTCACAC 1377  
Query 1316 CCGGATCAGCATGCTGCGGTGAATGCTTCCGAGTCTTGTACACACCGCCGTCACAC 1375  
1316 CCGGATCAGCATGCTGCGGTGAATGCTTCCGAGTCTTGTACACACCGCCGTCACAC 1375  
Query 1378 CATGGAGTGGGTTTCAAGAGAGGATGATGATGATGATGATGATGATGATGATGAT 1437  
1378 CATGGAGTGGGTTTCAAGAGAGGATGATGATGATGATGATGATGATGATGATGAT 1437  
Query 1376 CATGGAGTGGGTTTCAAGAGAGGATGATGATGATGATGATGATGATGATGATGAT 1435  
1376 CATGGAGTGGGTTTCAAGAGAGGATGATGATGATGATGATGATGATGATGATGAT 1435  
Query 1438 TGAGATTCAATGCTGGGGTG 1457  
1438 TGAGATTCAATGCTGGGGTG 1457  
Query 1436 TAGATTCAATGCTGGGGTG 1455  
1436 TAGATTCAATGCTGGGGTG 1455

RESULT 39  
US-10-029-397A-45  
Sequence 45, Application US/10029397A  
Publication No. US20030175709A1  
GENERAL INFORMATION:  
APPLICANT: MURPHY, GEORGE L.  
APPLICANT: WHITLEY, J. PENN.  
TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS  
FILE REFERENCE: AMB1:076US  
CURRENT APPLICATION NUMBER: US/10/029,397A  
NUMBER OF SEQ ID NOS: 73  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 45  
LENGTH: 1544  
TYPE: DNA  
ORGANISM: Neisseria meningitidis  
US-10-029-397A-45  
Query Match 76.9%; Score 1119.8; DB 6; Length 1544;



Best Local Similarity 87.1%; Pred. No. 1,7e-301;  
Matches 1275; Conservative 0; Mismatches 182; Indels 6; Gaps 4;

QY 1 ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAAACGACGAC--GGATGCTTGC 58  
DB 29 ATTGAACGCTGGCGGCGATGCTTTACACATGCAAGTCGAAACGACGACGACGACGACGCTTGC 88  
QY 59 ATCT--GGTGGCGAGTGGCGGAGCGGTGAGTAATGCATCGGAACGTAATCCAGAAAGGGG 116  
DB 89 TTCTCGGGTGGCGAGTGGCGAAGCGGTGAGTAATGCATCGGAACGTAATCCAGAGTGGG 148  
QY 117 GGTACGATCGCAAGATGCTTAATACCGCATATCTCTAAGAGAGAAAGCAGGGATC 176  
DB 149 GATTAATGATCGCAAGATGCTTAATACCGCATATCTCTTGAAGAGAAAGCAGGGATC 208  
QY 177 GAAGAAGCTTGGCGCTTTTGAAGCGGCGATGCTGATTAAGTGGTGGGTTAAAGGC 236  
DB 209 TTCCGGCGCTTGGCGCTTATTCGAGCGGCGCATATCTGATTAAGTGGTGGGTTAAAGGC 268  
QY 237 CTACCAAGGCGAGATCAGTATGTTGGTCTGAGAGAGACACACGCTTGGGATCTGAGA 296  
DB 269 CTACCAAGGCGAGATCAGTATGTTGGTCTGAGAGAGATGATCCGCCACACTGGGATGAGA 328  
QY 297 CACGGCCGAGACTCTTACGAGGAGAGCAGATGGGGAATTTTGAACAATGGCGCAAGCCT 356  
DB 329 CACGGCCGAGACTCTTACGAGGAGAGCAGATGGGGAATTTTGAACAATGGCGCAAGCCT 388  
QY 357 GATCCAGCAATGCGCGGTGAGTGAAGAAGCCTTGGGTTGTAAGCTCTTTCATGCGAG 416  
DB 389 GATCCAGCAATGCGCGGTGAGTGAAGAAGCCTTGGGTTGTAAGCTCTTTCATGCGAG 448  
QY 417 AAGAAAAGGTTACGTTAATATCTGACCCATGACCGGTATCGACAGAAAGACACCGGC 476  
DB 449 AAGAAAAGGCTGTGTTAATATCTGACCGGTATGACCGGTATCGAGAAATACACACCGGC 508  
QY 477 TAACTACGTCCAGCAGCGCGGTTAATAGTAGGAGTGAAGCGTTATCGGAATTTACTGG 536  
DB 509 TAACTACGTCCAGCAGCGCGGTTAATAGTAGGAGTGAAGCGTTATCGGAATTTACTGG 568  
QY 537 GCGTAAAGGAGTGGCAGCGCGCTTGAAGTCAATGTAAGTATCCCGGGCTTAACTGG 596  
DB 569 GCGTAAAGGAGTGGCAGCGCGCTTGAAGTCAATGTAAGTATCCCGGGCTTAACTGG 628  
QY 597 GAATTCGCTTTGAACTACAAAGCTTGAAGTGTGCGAGAGAGTGAATTCATGTGTA 656  
DB 629 GAATTCGCTTTGAACTACAAAGCTTGAAGTGTGCGAGAGAGTGAATTCATGTGTA 688  
QY 657 GCGAGTAAATGCGTAGATATGGAAGAACATGATGGCGAAGGCGAGCTCCCTGGGTTAA 716  
DB 689 GCGAGTAAATGCGTAGATATGGAAGAACATGATGGCGAAGGCGAGCTCCCTGGGTTAA 748  
QY 717 CACTACGCTCATGCAAGAAAGCGTGGGAGCAAGAGATTAGATACCTGGTAGTCCA 776  
DB 749 CACTACGCTCATGCAAGAAAGCGTGGGAGCAAGAGATTAGATACCTGGTAGTCCA 808  
QY 777 CCGCTTAAAGATGTCACTAGTTGTTGGCTTATTA--GGCTTGTAAACGAAGCTAACG 835  
DB 809 CCGCTTAAAGATGTCACTAGTTGTTGGCTTATTA--GGCTTGTGTAGAGCTAACG 868  
QY 836 CGTGAAGTTGACCGGCTGGGAGTACGCTGCAAGATTAACTCAAGAGAAATTGACGG 895  
DB 869 CGTGAAGTTGACCGGCTGGGAGTACGCTGCAAGATTAACTCAAGAGAAATTGACGG 928  
QY 896 GACCCGACAGAGCGGTGATATGATGATTAATTCAGTCAAGCGCAAAACCTTACCTA 955  
DB 929 GACCCGACAGAGCGGTGATATGATGATTAATTCAGTCAAGCGCAAAACCTTACCTA 988  
QY 956 CCCTTGACATGTAAGCAATTTTCTAGAGATTAATGTC--CTTGGGAACGCTAACACG 1014  
DB 989 GTCTTGACATGTAAGCAATTTTCTAGAGATTAATGTC--CTTGGGAACGCTAACACG 1048  
QY 1015 GTGCTGCATGCGTGTGTGACGCTCGTGTGTGAGATGTTGGTTAAGTCCCGAAGACG 1074

DB 1049 GTGCTGCATGCGTGTGTGACGCTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGC 1108  
QY 1075 GCAACCTTGTGATTAATTCGCAATGTTGGTGGGCACTTAAAGAGACTGGCGGTAC 1134  
DB 1109 GCAACCTTGTGATTAATTCGCAATGTTGGTGGGCACTTAAAGAGACTGGCGGTAC 1168  
QY 1135 AAACCGAGAAAGTGGGGAATGACGTCAGATCTTCATGAGCCCTTATGGGTAGGGCTTAC 1194  
DB 1169 AAACCGAGAAAGTGGGGAATGACGTCAGATCTTCATGAGCCCTTATGACAGGGCTTAC 1228  
QY 1195 ACGTAAATCAATGGCGCTACAGAGAGTGTCCCAACCGCGAGGGGAGCTTAATTCAGAA 1254  
DB 1229 ACGTAAATCAATGGCGCTACAGAGAGTGTCCCAACCGCGAGGGGAGCTTAATTCAGAA 1288  
QY 1255 ACGGCTGTGATCCGCGATCGGATCGGATCTGCAACTGCACTCCGTAAGTGGMAATGCTGT 1314  
DB 1289 AACCGATGTATCCGAGATGCACTCTCAACTGAGTCAAGTCAAGTGGMAATGCTGT 1348  
QY 1315 AATCGCGATCAGATGTCGCGTGAATACGTTCCCGGCTTGTACACACCGCCGTCA 1374  
DB 1349 AATCGCGATCAGATGTCGCGTGAATACGTTCCCGGCTTGTACACACCGCCGTCA 1408  
QY 1375 CACCATGGGAGTGGGTTTACCAAGAGAGTATGTTAATCGTAAAGAGGCGCTTGCCA 1434  
DB 1409 CACCATGGGAGTGGGGAATACCAAGAGTATGATTAACCAAGAGTCCGCTTACCA 1468  
QY 1435 CCGTGAATTCATGACTGGGCTG 1457  
DB 1469 CCGTATGCTTCAATGACTGGGCTG 1491

RESULT 40  
US-10-915-740A-1068  
; Sequence 1068, Application US/10915740A  
; Publication No. US20050191316a1  
; GENERAL INFORMATION:  
; APPLICANT: Frazer, Claire M.  
; APPLICANT: Hickey, Erin  
; APPLICANT: Peterson, Jeremy  
; APPLICANT: Tetteh, Heve  
; APPLICANT: Venter, J. Craig  
; APPLICANT: Masigiani, Vega  
; APPLICANT: Galeotti, Cesira  
; APPLICANT: Mora, Manosa  
; APPLICANT: Ratti, Giulio  
; APPLICANT: Scarfello, Maria  
; APPLICANT: Scarfello, Vincenzo  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizza, Mariagrazia  
; APPLICANT: Grandi, Guido  
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use  
; FILE REFERENCE: 002441.00090  
; CURRENT FILING DATE: US/10/915,740A  
; PRIOR FILING DATE: 2004-08-11  
; PRIOR APPLICATION NUMBER: 09/806,866  
; PRIOR FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: USSN 60/103,794  
; PRIOR FILING DATE: 1998-10-09  
; PRIOR APPLICATION NUMBER: USSN 60/132,068  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: PCT/US99/25373  
; PRIOR FILING DATE: 1999-10-08  
; NUMBER OF SEQ ID NOS: 1068  
; SOFTWARE: Patencin version 3.2  
; SEQ ID NO 1068  
; LENGTH: 2242716  
; TYPE: DNA  
; ORGANISM: Neisseria meningitidis  
US-10-915-740A-1068  
Query Match 76.9%; Score 1119.8; DB 9; Length 2242716;  
Best Local Similarity 87.1%; Pred. No. 9,7e-301;  
Matches 1275; Conservative 0; Mismatches 182; Indels 6; Gaps 4;





```
Db 1700902 ATTGAACGCTGGCGGATGCTTTACACATCCAAAGTCGAGCCGACGACAGAGAAAGCTTCC 1700843
QY 59 ATCT--GGTGGCGATGGCGGACCGGTGAGTAATGCATCGGAAAGTAATCCAGAAAGGGG 116
Db 1700842 TTCTCGGGGTGGCGAGATGGCGAAGCGGGTGAAGTAACATATCGGAAAGTAATCGAAGTAATGGGG 1700783
QY 117 GGTAAAGCATCGAAGAAATGTCATATACCGCATATCTCTAAAGGAAGAAAGCAAGGGGATC 176
Db 1700782 GATTAATGTAATCGAAGAAATGTAATACCGCATACCTCTTGAGAGAGAAAGCAAGGGGATC 1700723
QY 177 GAAAGACTTGGCGCTTTTGGAGCGCGCATGTCGTGATTAAGCTAGTGTGGGTAAAGGC 236
Db 1700722 TTGCGGCTTGGCGCTTATGAGGCGCGCATATCTGATTAAGCTAGTGTGGGTAAAGGC 1700663
QY 237 CTACCAAGGCGAGCATCAGTAGTGGTCTGAGAGAGACAGCAAGCCACACTGGGACTGAGA 296
Db 1700662 CTACCAAGGCGAGCATCAGTAGCGGGTCTGAGAGAGATGATCCGCCACACTGGGACTGAGA 1700603
QY 297 CAGGCGCCAGACTCCCTTAACGGGAGGCGACAGTGGGAAATTTTGGACATGGGCGCAAGCCT 356
Db 1700602 CAGGCGCCAGACTCTTAACGGGAGGCGACAGTGGGAAATTTTGGACATGGGCGCAAGCCT 1700543
QY 357 GATCCAGCAATGCGCGGTGAGTGAAGAAAGCCTTCGGGTTGTAAAGCTCTTTCAGTCGAG 416
Db 1700542 GATCCAGCCATGCGCGGTGCTGAAGAAAGCCTTCGGGTTGTAAAGCATTTTGTACAGG 1700483
QY 417 AAGAAAGGTTACGTTAATATCTGACCCATGACCGGTATGACAGAAAGACACCGCC 476
Db 1700482 AAGAAAGGCTGTTGCTAATATACGCGCGTGAAGACGTTAAGCTGAAGAAATTAAGCACCGCC 1700423
QY 477 TAACCTACGTCGACGACCGCGGTATATAGTAGGGTGAAGCGCTTAATCGGAATTCGAG 536
Db 1700422 TAACCTACGTCGACGACCGCGGTATATAGTAGGGTGAAGCGCTTAATCGGAATTCGAG 1700363
QY 537 GCGTAAAGGGTGGCGAGCGGCGCTTGAAGTCAGATGTGAATCCCGGGCTTAACTACCTG 596
Db 1700362 GCGTAAAGGGGCGGCGAGCGGTACTTAAGCAGAGATGTGAATCCCGGGCTTAACTACCTG 1700303
QY 597 GAATTCGCTTTGAACCTACAAAGCTAAGAGTGGCGAGAGGAGTGAATTCATGTGTA 656
Db 1700302 GAATTCGCTTTGAACCTGAGGTGACTGAGTGTCTCAGAGGAGGTGAATTCACCTGTA 1700243
QY 657 GAGTGAATATGCTTAAGATATGGAAGAAACATGAGAGGCGAAGCGCTTCCTGGGTAA 716
Db 1700242 GAGTGAATATGCTTAAGATATGGAAGAAACATGAGAGGCGAAGCGCTTCCTGGGTAA 1700183
QY 717 CACTGACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAATCCTGTAGTACCA 776
Db 1700182 CACTGACGCTCATGACCGGAAAGCGTGGGAGCAAAACAGATTAATCCTGTAGTACCA 1700123
QY 777 CGGCTTAACGATGCACTAGTGTGGGCTTATTA--GGCTTGTAACGAAGCTAACG 835
Db 1700122 CGGCTTAACGATGCACTAGTGTGGGCAACTGATTTGCTGTAGCGTACGTAACG 1700063
QY 836 CGTGAAGTTGACCGCTGGGGAGTACGTCGCAAGTTTAAATTCGAAAGAAATTTGACGG 895
Db 1700062 CGTGAAGTTGACCGCTGGGGAGTACGTCGCAAGTTTAAATTCGAAAGAAATTTGACGG 1700003
QY 896 GACCCGCAAGAGGCGTGAATTAATGATTAATTCGATGCAAGCGGAAAGAACTTACCTA 955
Db 1700002 GACCCGCAAGAGGCGTGAATTAATGATTAATTCGATGCAAGCGGAAAGAACTTACCTA 1699943
QY 956 CCGTTGACATGTAGCGAATTTTCTAGAGATAGATTAGTG--CTTCGGGAAACGCTAACAG 1014
Db 1699942 GCTTGAACATGTAGCGAATTCCTCCGAGACGAGAGAGTGCTTCGGGAGCGGTAAACAG 1699883
QY 1015 GTGCTGCAATGGCTGTCTCAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGAAGAAC 1074
Db 1699882 GTGCTGCAATGGCTGTCTCAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGAAGAAC 1699823
QY 1075 GCAACCTTGTCTTAATTTGATGATGCTTTGGTGGGCACTTAAATGAGCTGCGGTGAC 1134
```

```
Db 1699822 GCAACCTTGTCTTAATTTGATGATTCATCATTCAGTTGGGCACTTAATGAGATGCGCGTAC 1699763
QY 1135 AACCAGGAGAGTGGGGATGACGTCAAGTCTCATGCGCCTTAATGGGTAGGGCTTAC 1194
Db 1699762 AAGCCGAGAGAGTGGGGATGACGTCAAGTCTCATGCGCCTTAATGAGATGCGCGCTTAC 1699703
QY 1195 ACGTAATCAATGAGCGGTACAGAGGGTGGCCAAACCCGAGAGGGGAGTAAATTCAGAA 1254
Db 1699702 ACGTATCAATATGTCGCGTACAGAGGGTAAACCAACCCGAGAGGGGAGCAATTCACAA 1699643
QY 1255 ACGCGTGTAGTCCGATCGGATCGAGTCTGCACTCCGTAAGTGGAAATCGCTAGT 1314
Db 1699642 AACCGATGTAGTCCGATCGGATCGAGTCTGCACTCCGTAAGTGGAAATCGCTAGT 1699583
QY 1315 AATCGCGATGACATGTCGCGGTAAATACGTTCCCGGGTCTTGAACAACCGCGCGTCA 1374
Db 1699582 AATCGCAGGTGACATGCTGCGGTAAATACGTTCCCGGGTCTTGAACAACCGCGCGTCA 1699523
QY 1375 CACCATGGGAGTGGGTTTCAACGAGAGAGTGTCTTAACCGTAAAGAGGCGCTTGCCA 1434
Db 1699522 CACCATGGGAGTGGGAGTACCAAGAGTAAAGTAAAGATTAACCAAGAGATCGCTTACCA 1699463
QY 1435 CGGTGAGATTCAATGACTGGGTTG 1457
Db 1699462 CGGTATGCTTCAATGACTGGGTTG 1699440

RESULT 42
US-10-029-397A-43
; Sequence 43, Application US/10029397A
; Publication No. US20030175709A1
; GENERAL INFORMATION:
; APPLICANT: WHITLEY, J. PENN
; TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING RNA POPULATIONS
; FILE REFERENCE: AMB1:076US
; CURRENT APPLICATION NUMBER: US/10/029,397A
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patemlin Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1610
; TYPE: DNA
; ORGANISM: Burkholderia pseudomallei
US-10-029-397A-43

Query Match 76.8%; Score 118.8; DB 6; Length 1610;
Best Local Similarity 86.8%; Pred. No. 3.3e-301;
Matches 1268; Mismatches 182; Indels 10; Gaps 3;

QY 1 ATTGAACCTGGCGGATGCTTTACACATCCAAAGTCGAGCCGACGACAGAGAAAGCTTCCAT 60
Db 82 ATTGAACCTGGCGGATGCTTTACACATCCAAAGTCGAGCCGACGACGAG--GCTTCGCG 139
QY 61 CTGATGCGAGTGGCGGACCGGTGATGATCGAAGCTATCCAGAAAGGGGGGTA 120
Db 140 CTGATGCGAGTGGCGGACCGGTGATGATCGAAGCTATCCAGAAAGGGGGGTA 199
QY 121 ACGCATCGAAGATGCTAATACCGCATATATCTTAAGAGAGAAAGCAAGGGGATCGAAA 180
Db 200 GCCCGCGAAGACCGAATTAATACCGCATATATCTTAAGAGAGAAAGCAAGGGGATCGAAA 259
QY 181 GACCTTGGGCTTTTGAAGCGCGCATGTCTGATTAAGTGTGGTGGGTAAAGGCTTAC 240
Db 260 GGCCTCGGCTATAGGGTGGCGCATGTCTGATTAAGTGTGGTGGGTAAAGGCTTAC 319
QY 241 CAAGGCGAGATCAGTATGTTGCTGAGAGGACGACAGCCACACTGGGACTGAGACAG 300
Db 320 CAAGGCGAGATCAGTATGTTGCTGAGAGGACGACAGCCACACTGGGACTGAGACAG 379
QY 301 GCCCAGACTCTTACCGGAGGACGAGTGGGGAATTTTGAACATAGGCGCAAGCTGATC 360
Db 380 GCCCAGACTCTTACCGGAGGACGAGTGGGGAATTTTGAACATAGGCGCAAGCTGATC 439
```

Oy	36	CAGCAATGCCGCGAGGAGGAAAGAGGCTTCGGGTTGTAAAGCTCTTCAAGTGAAGA	420
Db	440	CAGCAATGCCGCGGTGTGAAAGAGGCTTCGGGTTGTAAAGCACTTTGTCCGAAAGA	439
Oy	421	AAAGTTACGGTAATAATACTGTGCCATCGATGACGATTCGACAGAGAAAGCACCGGCTAAC	480
Db	500	AATGATCTTGGCTAATATCCGGAGTGGATGACGGTACCCGAAAGATTAAGCACCGGCTAAC	559
Oy	481	TACGTGCACAGCCCGCGGTAAATCGTAGGGTGCAGCCTTAATCGAATTAATCTGGGCGT	540
Db	560	TACGTGCACAGCCCGCGGTAAATCGTAGGGTGCAGCCTTAATCGGAAATTAATCTGGGCGT	619
Oy	541	AAAGGTGCCGAGCGCGCTTGTAAATCGATGGAATGGAATCCCGGGGCTTAACCTGGGAAT	600
Db	620	AAAGCGTGCAGGCGGTTGTCTAAGACCGATGGAATCCCGGGGCTTAACCTGGGAAC	679
Oy	601	TGCGTTTGAATCTCAAAAGCTAGAGTGTGCGACGAGGAGGTGAAATTCATGTGTACAG	660
Db	680	TGCATTTGTGACTGGCAGGCTAGAGTATGCGAGAGGGGGGTAGAAATTCACGTGTACAG	739
Oy	661	TGAATGCGTAGAGATATGGAAGAACATCGATGCGAAGGCAAGCTTCCTGGGTTAACACT	720
Db	740	TGAATGCGTAGAGATATGGAAGAACATCGATGCGAAGGCAAGCTTCCTGGGCTAACACT	739
Oy	721	GACGCTACGACGAAAGCGTGGGAGCAACAGGATTAATACCTCGTGTATGTCAAGCC	780
Db	800	GACGCTACGACGAAAGCGTGGGAGCAAAACAGGATTAATACCTCGTGTATGTCAAGCC	859
Oy	781	CTAAACGATGCATCTAGTTGTGGGCGCTTATGAGGCTGGTATCGAAGCTAACGCGTGA	840
Db	860	CTAAACGATGCATCTAGTTGTGGGCGCTTATGAGGCTGGTATCGAAGCTAACGCGCGA	919
Oy	841	AGTTGACCGCTGGGAGTAGACGATCGCAAGATTAAACTCAAGAAATTGACGCGGAGCC	900
Db	920	AGTTGACCGCTGGGAGTAGACGATCGCAAGATTAAACTCAAGAAATTGACGCGGAGCC	979
Oy	901	GCACAAAGCGGTGATTAATGTGATTAATTCATGCAACGCAAAACCTTAATCTAACCTT	960
Db	980	GCACAAAGCGGTGATTAATGTGATTAATTCATGCAACGCAAAACCTTAATCTAACCTT	1039
Oy	961	GACATGTAGCGAATTTTCTAGAGATAGATTAGTGT---TCGGGAGCGCTAACAGAGTG	1017
Db	1040	GACATGTGTGGAAAGCCCGATAGAGATTGGGCGTGTGCAAAAGAAACCGCGCACAGTGT	1099
Oy	1018	CTGCATGGCTGTGCTCAGCTCGTGTGCGAGATGTGTGGTTAAGTCCCGCAACAGCGCA	1077
Db	1100	CTGCATGGCTGTGCTCAGCTCGTGTGCGAGATGTGTGGTTAAGTCCCGCAACAGCGCA	1159
Oy	1078	ACCCTTGTCAATTAATTTGGCATCATTTGGTTGGGCACTTTAATGAGACTGCGGTGACAA	1137
Db	1160	ACCCTTGTCTTAATTTGTCTAC----GCAGAGGCACTTAAAGAGACTGCGGTGACAA	1214
Oy	1138	CCGAGAGAAAGTGGGAGTAGCTCAAGTCTCAATGACCCTTAATGAGGTAGGCTTCAACG	1197
Db	1215	CCGAGAGAAAGTGGGAGTAGCTCAAGTCTCAATGACCCTTAATGAGGTAGGCTTCAACG	1274
Oy	1198	TAAATCAATGGCGCTACAGAGGGTTGCGCAACCGCGAGGGGAGCTAATCTCAGAAAGC	1257
Db	1275	TCATACAAATGTGCGGAACAGAGGGTGTGCGCAACCGCGAGGGGAGCCAAATCCAGAAAC	1334
Oy	1258	GCGTGTAGTCCGGAATCGGAATCTGCAACTGCACTCCGTGAAGTTCGAAATCGCTAGTAA	1317
Db	1335	CGATGTAGTCCGGAATCTGCAACTGCAACTCGAATGCAATGGAAGTGGAAATCGCTAGTAA	1394
Oy	1318	CGCGAATCAGATGTGCGCGTGAATACGTTCCCGGCTTTGTACACACCGCCCGTCAAC	1377
Db	1395	CGCGAATCAGATGTGCGCGTGAATACGTTCCCGGCTTTGTACACACCGCCCGTCAAC	1454
Oy	1378	CATGGAGTGGGTTTACACAGAGAGGATGCTAACCGTAAAGAGGGGCGCTGTGCACGG	1437
Db	1455	CATGGAGTGGGTTTACACAGAGAGGCTAGTCTAACCGCAAGAGAGCGGTCAACACAGG	1514

Oy	1438	TGAGATTCAATGACTGGGGTG	1457
Db	1515	TAGAATTCATGACTGGGGTG	1534
RESULT 43			
US-09-726-774-5			
; Sequence 5, Application US/09726774			
; Patent No. US2002008226A1			
; GENERAL INFORMATION:			
; APPLICANT: Iversen, Patrick L.			
; TITLE OF INVENTION: Antisense Antibacterial Method and			
; TITLE OF INVENTION: Composition			
; FILE REFERENCE: 0450-0032.30			
; CURRENT APPLICATION NUMBER: US/09/726,774			
; CURRENT FILING DATE: 2000-11-29			
; PRIOR APPLICATION NUMBER: US 60/168,150			
; PRIOR FILING DATE: 1999-11-29			
; NUMBER OF SEQ ID NOS: 139			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; LENGTH: 1544			
; TYPE: DNA			
; ORGANISM: Neisseria gonorrhoea			
US-09-726-774-5			
Query Match                76.7%; Score 118.2; DB 3; Length 1544;			
Best Local Similarity     87.1%; Pred.No.4.7e-30;			
Matches 1274; Conservative 0; Mismatches 183; Indels 6; Gaps 4			
Oy	1	AATTGAACCTTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGCAC--GGATGCTTGC	58
Db	29	ATTGAACCTTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGCACGGAACTTGC	88
Oy	59	ATCT--GGTGGAGATGTGCGGACGGGGTAGTAATCATTCGNAACGTATCCAGAAAGGGG	116
Db	89	TTCTGGGTGGCGAATGTGCGGACGGGGTAGTAATCATTCGNAACGTATCCGGTAGCGGG	148
Oy	117	GATAACGATCGAAAGATGTGCTAATACCGCATATATCTTAAAGAGAAAGCAGGGGATC	176
Db	149	GATACTGATCGAAAGATCAGCTAATATCCGATATGCTTGAAGAGGAAAGCAGGGGACC	208
Oy	177	GAAGAACCCTTGGGCTTTGGAGCCGCCCATCTGATTAGCATGTGGTGGGTAAAGGC	236
Db	209	TTCCGGGCTTTGGGCTATCCGAGCGGCCCATATCTGATTAGCTGTGGGGGTAAAGGC	268
Oy	237	CTAACGAGCGCATCTCTACCGGAGCGAGCAGTAGTGGGAATTTTGAACAATGGCGCAAGCCT	296
Db	269	CCACCAAGCGCATCTCTACCGGAGCGAGCAGTAGTGGGAATTTTGAACAATGGCGCAAGCCT	328
Oy	297	CACGGCCGAGACTCTCTACCGGAGCGAGCAGTAGTGGGAATTTTGAACAATGGCGCAAGCCT	356
Db	329	CACGGCCGAGACTCTCTACCGGAGCGAGCAGTAGTGGGAATTTTGAACAATGGCGCAAGCCT	388
Oy	357	GATTCAGCATATGCCCGTAGTGAAGAAAGCCTTGGGTTTGAAGCTCTTCAGTCGAG	416
Db	389	GATTCAGCATATGCCCGTAGTGAAGAAAGCCTTGGGTTTGAAGCTCTTCAGTCGAG	448
Oy	417	AAGAAAAGGTTACGGTAAATATATGTCGACCCCATGACGGATCGACAGAAAGACACCGGC	476
Db	449	AAGAAAAGGCTTGTCCAAATATCGGGGGCGATGACGGATCTTGAAGAAATTAAGACACCGGC	508
Oy	477	TAACTACGTCGACAGAGCCGCGGTATATAGTAGGGTGCAGAGCGTTAATCGAATTACTGG	536
Db	509	TAACTACGTCGACAGAGCCGCGGTATATAGTAGGGTGCAGAGCGTTAATCGAATTACTGG	568
Oy	537	GGCTAAAGGGTGCAGAGCGCGCTTTGTATGATCATATGTGAATCCCCCGGCTTAACTGG	596
Db	569	GGCTAAAGGGTGCAGAGCGGTATCTTAAGCAGATGTGAATCCCCCGGCTTAACTGG	628
Oy	597	GAAATGCGTTTGAATCAAAAGCTAAGATGTGGAGAGAGGAGTGGAAATTCATATGTGA	656
Db	629	GAACTGCGTTTGAATCGAGTGTGTGATGAGAGAGTGGAAATTCATATGTGA	688

```
QY 657 GCAGTGAATGCTAGATATGGAAGAACTGATGGCAAGGCAAGCTCTGGGTAA 716
DB 689 GCAGTGAATGCTAGATATGGAAGAACTGATGGCAAGGCAAGCTCTGGGTAA 748
QY 717 CACTGACGCTCATGCAAGAAAGCGTGGGAGCAAAAGATTAATGATACCTGGTATGCA 776
DB 749 CACTGACGCTCATGCTCGAAAGCGTGGGAGCAAAAGATTAATGATACCTGGTATGCA 808
QY 777 CGCCCTAAACGATGTCATAGTGTGTTGGCTTATTA -GGCTTGTAAAGAACTAACG 835
DB 809 CGCCCTAAACGATGTCATAGTGTGTTGGCACTTATGCTTGTAGGTAGCTAACG 868
QY 836 CCGTAAGTTGACCGCTCGGGAGATACGCTGCGAAAGTTAAACTCAAGAAATTAACGCG 895
DB 869 CGTAAGTTGACCGCTCGGGAGATACGCTGCGAAAGTTAAACTCAAGAAATTAACGCG 928
QY 896 GACCCGCAAGCGGCTGATATGATGATTAATTCATGCAAGCGCAAAACCTTAACCTA 955
DB 929 GACCCGCAAGCGGCTGATGATGATTAATTCATGCAAGCGCAAAACCTTAACCTA 988
QY 956 CCCTTGACATGATGCAATTTTCTAGAGATATGATGATG -CTTCCGGAACGCTAACAC 1014
DB 989 GTTTGACATGATGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATG 1048
QY 1015 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1074
DB 1049 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1108
QY 1075 GCAACCTTGTCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1134
DB 1109 GCAACCTTGTCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1168
QY 1135 AAACCGGAGAGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 1194
DB 1169 AAGCCGAGAGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1228
QY 1195 AGCTAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1254
DB 1229 AGCTAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1288
QY 1255 AGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1314
DB 1289 AACCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1348
QY 1315 AATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1374
DB 1349 AATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1408
QY 1375 CACCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1434
DB 1409 CACCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1468
QY 1435 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1457
DB 1469 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1491
```

```
RESULT 44
US-10-029-397A-44
; Sequence 44, Application US/10029397A
; Publication No. US20030175709A1
; GENERAL INFORMATION:
; APPLICANT: MURPHY, GEORGE L.
; APPLICANT: WHITLEY, J. PENN
; TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING RNA POPULATIONS
; FILE REFERENCE: AMB1:076US
; CURRENT APPLICATION NUMBER: US/10/029,397A
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 1544
```

```
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-029-397A-44
Query Match 76.7%; Score 1118.2; DB 6; Length 1544;
Best Local Similarity 87.1%; Pred. No. 4,7e-301;
Matches 1274; Conservative 0; Mismatches 183; Indels 6; Gaps 4;
QY 1 ATTGAACGCTGGCGGATGCTTATACATGCAAGTCGAACGGCAGAC--GGATGCTTGC 58
DB 29 ATTGAACGCTGGCGGATGCTTATACATGCAAGTCGAACGGCAGACGGAACTTGC 88
QY 59 ATCT--GGTGGGATGCGGAGCGAGTATGATGATGATGATGATGATGATGATGATGATG 116
DB 89 TTCTGGGATGCGGAGCGGAGTATGATGATGATGATGATGATGATGATGATGATGATG 148
QY 117 GGTAAAGCATGAAAGATGCTTAATACCGATATATCTTAAGAGAAAGCAAGGAGATC 176
DB 149 GATTAAGTATGAAAGATGCTTAATACCGATATATCTTAAGAGAGAAAGCAAGGAGATC 208
QY 177 GAAAGACCTTGGGCTTTTGGAGCGGCGATGCTGATTAATGATGATGATGATGATGATG 236
DB 209 TTTCGGGCTTGGGCTTATCCGATGCGGCGATATCTGATTAATGATGATGATGATGATG 268
QY 237 CTACCAAGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 296
DB 269 CCACCAAGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 328
QY 297 CACGCGCCAGACTCTTACCGGAGGAGCAGATGAGGAAATTTTGGACATGAGGCGCAAGCT 356
DB 329 CACGCGCCAGACTCTTACCGGAGGAGCAGATGAGGAAATTTTGGACATGAGGCGCAAGCT 388
QY 357 GATTCAGCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 416
DB 389 GATTCAGCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 448
QY 417 AAGAAAGGTTACGCTTAATATCGTGAACCATGACGATGATGATGATGATGATGATGATG 476
DB 449 AAGAAAGGCTTGTGCAATATCGGCGGCGATGATGATGATGATGATGATGATGATGATG 508
QY 477 TAACTACGTGCAAGACCGCGGATATACGATGATGATGATGATGATGATGATGATGATG 536
DB 509 TAACTACGTGCAAGACCGCGGATATACGATGATGATGATGATGATGATGATGATGATG 568
QY 537 GGTAAAGGCTGCGGAGCGGCTTGTATGATGATGATGATGATGATGATGATGATGATG 596
DB 569 GGTAAAGGCTGCGGAGCGGCTTGTATGATGATGATGATGATGATGATGATGATGATG 628
QY 597 GAATGCGTTTGAATTAACAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 656
DB 629 GAATGCGTTTGAATTAACAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 688
QY 657 GCAGTGAATGCTAGATATGGAAGAACTGATGGCAAGGCAAGCTCTGGGTAA 716
DB 689 GCAGTGAATGCTAGATATGGAAGAACTGATGGCAAGGCAAGCTCTGGGTAA 748
QY 717 CACTGACGCTCATGCAAGAAAGCGTGGGAGCAAAAGATTAATGATACCTGGTATGCA 776
DB 749 CACTGACGCTCATGCTCGAAAGCGTGGGAGCAAAAGATTAATGATACCTGGTATGCA 808
QY 777 CGCCCTAAACGATGTCATAGTGTGTTGGCTTATTA -GGCTTGTAAAGAACTAACG 835
DB 809 CGCCCTAAACGATGTCATAGTGTGTTGGCACTTATGCTTGTAGGTAGCTAACG 868
QY 836 CCGTAAGTTGACCGCTCGGGAGATACGCTGCGAAAGTTAAACTCAAGAAATTAACGCG 895
DB 869 CGTAAGTTGACCGCTCGGGAGATACGCTGCGAAAGTTAAACTCAAGAAATTAACGCG 928
QY 896 GACCCGCAAGCGGCTGATATGATGATTAATTCATGCAAGCGCAAAACCTTAACCTA 955
DB 929 GACCCGCAAGCGGCTGATGATGATTAATTCATGCAAGCGCAAAACCTTAACCTA 988
QY 956 CCCTTGACATGATGCAATTTTCTAGAGATATGATGATG -CTTCCGGAACGCTAACAC 1014
```

Db 989 GTTTGACATGTCGGAAATCCCTCCGAGACGAGAGAGTGCCTTCGAGACCCTTAACACAG 1048  
 Qy 1015 GTGCTGCATGCGTGTGTCAGCTCGTGTGAGAGTGTGGTAAAGTCCGCAACAGC 1074  
 Db 1049 GTGCTGCATGCGTGTGTCAGCTCGTGTGAGAGTGTGGTAAAGTCCGCAACAGC 1108  
 Qy 1075 GCAACCTTGTCTAATTAATTCGATCTTGTGGGCACTTTAATGAGACTCCGGTGC 1134  
 Db 1109 GCAACCTTGTCTAATTAATTCGATCTTGTGGGCACTTTAATGAGACTCCGGTGC 1168  
 Qy 1135 AAACCGAGAGAGTGGGAGTACGTCAAGTCTTATGAGCTTATGGTACGCTTAC 1194  
 Db 1169 AAGCCGAGAGAGTGGGAGTACGTCAAGTCTTATGAGCTTATGGTACGCTTAC 1228  
 Qy 1195 AGTAATTAATTAATGAGGCGTACAGAGGAGTGGCAACCCGAGAGGAGTAACTCAG 1254  
 Db 1229 AGTCAATCAATGTCGCGTACAGAGGAGTGGCAACCCGAGAGGAGTAACTCAG 1288  
 Qy 1255 AGCGGCTGTAATGCGGATCGAGTCTGCACTCGAATCGGAAATCGGTAAT 1314  
 Db 1289 AACCATGCTAATCGCGAATTCGACTCTGCACTCGAATCGGAAATCGGTAAT 1348  
 Qy 1315 AATCGCGATCAAGATGTCGCGTAAATCGTTCGCGGCTTTGTAACAACCGCCGCT 1374  
 Db 1349 AATCGCGATCAAGATGTCGCGTAAATCGTTCGCGGCTTTGTAACAACCGCCGCT 1408  
 Qy 1375 CACCATGAGAGTGGGTTTCAACGAGAGTGTAAACGTAACGTAAGAGAGCGCTTGC 1434  
 Db 1409 CACCATGAGAGTGGGTTTCAACGAGAGTGTAAACGTAACGTAAGAGAGCGCTTGC 1468  
 Qy 1435 CGGTGATTCATGACTGGGGTG 1457  
 Db 1469 CGGTGATTCATGACTGGGGTG 1491

RESULT 45  
 US-10-719-633-5  
 ; Sequence 5: Application US/10719633  
 ; Publication No. US20040137485A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Iversen, Patrick L.  
 ; TITLE OF INVENTION: Antisense Antibacterial Method and  
 ; FILE REFERENCE: 0450-0032.30  
 ; CURRENT APPLICATION NUMBER: US/10/719, 633  
 ; CURRENT FILING DATE: 2003-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/726,774  
 ; PRIOR FILING DATE: 2000-11-29  
 ; PRIOR APPLICATION NUMBER: US 60/168,150  
 ; PRIOR FILING DATE: 1999-11-29  
 ; NUMBER OF SEQ ID NOS: 139  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 1544  
 ; TYPE: DNA  
 ; ORGANISM: *Neisseria gonorrhoea*  
 US-10-719-633-5

Query Match 76.7%; Score 1118.2; DB 7; Length 1544;  
 Best Local Similarity 87.1%; Pred. No. 4,7e-301;  
 Matches 1274; Conservative 0; Mismatches 183; Indels 6; Gaps 4;  
 Qy 1 ATTTGAAGCTGCGGCGATGCTTTTACATGCAAGTCAAGCGACGAC--GGATGCTTGC 58  
 Db 29 ATTTGAAGCTGCGGCGATGCTTTTACATGCAAGTCAAGCGACGACGAGAAAGCTTGC 88  
 Qy 59 ATCT--GGTGGAGAGTGGCGAGCGGTGAGTAAATGATCGAAAGTATCCAGAGAGGG 116  
 Db 89 TTCTCGGATGGCGAGTGGCGAGCGGTGAGTAAATGATCGAAAGTATCCAGAGAGGG 148  
 Qy 117 GGTAAAGCATGAAAGATGTCTAATATCCGATATACTCTTAAGAGAGAAAGCGGGATC 176

Db 149 GATAACTGATCGAAGATACAGCTAATACCGCATACGTCCTTGAGAGGAGAAACAGGGAGC 208  
 Qy 177 GAAAGACTTGTGCGCTTTTGGAGCGCGCGATGTCTGATTAAGTGTGGGTAAAGGC 236  
 Db 209 TTGCGGCTTGCCTATCCGACCGCCGATATCTGATTAAGTGTGGGTAAAGGC 268  
 Qy 237 CTACCAAGCGACAGTACGATGTGTGTGAGAGACACAGCCACATGAGACTGAG 296  
 Db 269 CCACCAAGCGACAGTACGATGTGTGTGAGAGATATCCGACATGAGACTGAG 328  
 Qy 297 CACGCGCCAGACTCTTACGAGAGGAGCAGTGGGAAATTTTGAACAATGGCGCAAGCCT 356  
 Db 329 CACGCGCCAGACTCTTACGAGAGGAGCAGTGGGAAATTTTGAACAATGGCGCAAGCCT 388  
 Qy 357 GATCCAGCAATGCGCGTGAATGAAGAGCTTTCGGTTTGAAGCTCTTCACTCAG 416  
 Db 389 GATCCAGCAATGCGCGTGAATGAAGAGCTTTCGGTTTGAAGCTCTTCACTCAG 448  
 Qy 417 AAGAAAGGTTACGGTAAATATCGTGACCCATGACGGTATCGACAGAGAACCCGAC 476  
 Db 449 AAGAAAGGCTGTTCCAAATATCGCGCGCCGATGACGTTACTGAAGAAATAGCACCGGC 508  
 Qy 477 TAACTACGTGCGAGAGCGCGGTAAATCGTAGGTTGCAAGCTTAATCGAATTAATCTGG 536  
 Db 509 TAACTACGTGCGAGAGCGCGGTAAATCGTAGGTTGCAAGCTTAATCGAATTAATCTGG 568  
 Qy 537 GGTAAAGGTTGCGAGAGCGCGCTTGTAAATGATGTAATATCCCGGGCTTAATCTGG 596  
 Db 569 GGTAAAGGTTGCGAGAGCGCGCTTGTAAATGATGTAATATCCCGGGCTTAATCTGG 628  
 Qy 597 GAAATTCGCTTAAACTCAAAAGCTAAGATGTCGAGAGGAGTGAATTCATGTGT 656  
 Db 629 GAAATTCGCTTAAACTGAGTGAATGTCGAGAGGAGTGAATTCATGTGT 688  
 Qy 657 GCAAGTAAATGCGTGAATGATGTAAGAAATCGATGAGAGAGCAGCTCTGGTTAA 716  
 Db 689 GCAAGTAAATGCGTGAATGATGTAAGAAATCGATGAGAGAGCAGCTCTGGTTAA 748  
 Qy 717 CACTGACGTCATGACGAAAGCGTGGGAGGAAACAGGATTAATGATCCCTGGTATCCA 776  
 Db 749 CACTGACGTCATGACGAAAGCGTGGGAGGAAACAGGATTAATGATCCCTGGTATCCA 808  
 Qy 777 CGCCCTAAACGATGCAACTAGTGTGGGCTTATTA--GGCTGTAACGAAGCTTAAG 835  
 Db 809 CGCCCTAAACGATGCAACTAGTGTGGGCTTATTA--GGCTGTAACGAAGCTTAAG 868  
 Qy 836 CGTGAAGTGAACCGCTGGGAGTACCGTCCGAAGTTAAATCTCAAGAAATGACGGG 895  
 Db 869 CGTGAAGTGAACCGCTGGGAGTACCGTCCGAAGTTAAATCTCAAGAAATGACGGG 928  
 Qy 896 GACCGCGACAGCGGTGATTAATGATGTAATTCGATGCAACCGGAAACCTTACTTA 955  
 Db 929 GACCGCGACAGCGGTGATTAATGATGTAATTCGATGCAACCGGAAACCTTACTTA 988  
 Qy 956 CCTTGAACATGAGGAAATTTTCTAGAGTAAATTAAGT--CTTCGGGAAAGCGTAAACAG 1014  
 Db 989 GTTTGACATGTGCGGAAATCTCCGAGACGAGAGGATCTTCGGGAAAGCGTAAACAG 1048  
 Qy 1015 GTGCTGCATGCGTGTGTCAGCTCGTGTGAGAGTGTGGTAAAGTCCGCAACAGC 1074  
 Db 1049 GTGCTGCATGCGTGTGTCAGCTCGTGTGAGAGTGTGGTAAAGTCCGCAACAGC 1108  
 Qy 1075 GCAACCTTGTCTAATTAATTCGATCTTGTGGGCACTTTAATGAGACTCCGGTGC 1134  
 Db 1109 GCAACCTTGTCTAATTAATTCGATCTTGTGGGCACTTTAATGAGACTCCGGTGC 1168  
 Qy 1135 AAACCGAGAGAGTGGGAGTACGTCAAGTCTTATGAGCTTATGGTACGCTTAC 1194  
 Db 1169 AAACCGAGAGAGTGGGAGTACGTCAAGTCTTATGAGCTTATGGTACGCGGCTTAC 1228  
 Qy 1195 AGTAATTAATTAATGAGGCGTACAGAGGAGTGGCAACCCGAGAGGAGTAACTCAG 1254  
 Db 1229 AGTAATTAATTAATGAGGCGTACAGAGGAGTGGCAACCCGAGAGGAGTAACTCAG 1288

[illegible]

QY	431	GTAAATTAATCGGACCCCAATGACCGATCGACAGAAAGACACCGGCTTAATCTACGTCGAG	490
Db	431	GTTAATTAATCTCGGGGGGAGTAGACCGTATCCGGAAAAATTAAGCAACCGGCTTAATCTACTTGCCA-	489
QY	491	CAGCCGCGGTAAATACGTAGAGGGGTGCAGACGCTTAATCGGAATTACTGGGCTTAAAGGCTGC	550
Db	490	CAGCCGCGGTAAATACGTAGAGGGGTGCAGACGCTTAATCGGAATTACTGGGCTTAAAGGCTGC	549
QY	551	CAGGGGCGCTTGTAAGTACAGATGTGAATCCCGGGCTTAACCTGGGAATTTGGCTTGA	610
Db	550	CAGGCGGTTTTGTATAAGACGGAATGTAAATCCCCGGCTTAACCTGGGAATTTGGCTTGA	609
QY	611	ACTACAAAGCTAGAGTGTGCGCAGAGGGAGGTGGAATTCATGTGTAGCAGTGAATCGCT	670
Db	610	ACTGGAAGGCTAGAGTATGCGCAGAGGGGGGTAGAAATTCACGTTATGACGTGAATAGCGT	669
QY	671	AGAGATTAATGAAAGAAATCATGATGCGGAAGCGACCTCCTGGGTTAACTAGACGCTCATG	730
Db	670	AGAGATGTGAGGAATTAACGATGTGCGGAAGCGCCCTCGGGCCAAATATCTGACGCTCATG	729
QY	731	CACGGAACCGTCGGGAGCAAAACAGAAATTAGATACCCTGTATGCCACGCGCTTAAGAGATG	790
Db	730	CACGGAACCGTCGGGAGCAAAACAGAAATTAGATACCCTGTATGCCACGCGCTTAAGAGATG	789
QY	791	TCAAATTAATTTGTTGGGCGCTTAATTAAGGCTTGTGTAAAGAACTAACGCGTGAAGTTGACCG	850
Db	790	TCAAATTAATTTGTTGGGAGATTAATTTCTTCTTAATGAAGAACTAACGCGTGAAGTTGACCG	849
QY	851	CTGGGAGAGTACCGTTCGCAAGATTAAAACTCAAGAAATTGACGGGGAACCCGCAACAGCGG	910
Db	850	CTGGGAGAGTACCGTTCGCAAGATTAAAACTCAAGAAATTGACGGGGAACCCGCAACAGCGG	909
QY	911	TGGATTAATGTGGATTAATTTGCATGGAACGCGAAAAACCTTAACCTACCTTTGACATGTAGC	970
Db	910	TGGATTAATGTGGATTAATTTGCATGGAACGCGAAAAACCTTAACCTACCTTTGACATGTAGC	969
QY	971	GAATTTTCTAGAGATAGATTAGTGTCT--TCCGGAACGCTAACACAGGTGCTGATGAGCT	102
Db	970	GAATTTTCTAGAGATAGATTAGTGTCTCGAAGAAAGAACGTTAACACAGGTGCTGATGAGCT	102
QY	1028	GTCGTACAGCTCGTGTTCGTGAGATGTTGGGTTAAATGCCGCAACGAGCGCAACCTTTGTCA	108
Db	1030	GTCGTACAGCTCGTGTTCGTGAGATGTTGGGTTAAATGCCGCAACGAGCGCAACCTTTGTCC	108
QY	1088	TTAATTTGCATTTGGTGTGGGCACTTAATGAGACTGCGGATGCAAAACGAGAGGAAG	114
Db	1090	TTAATTTGC-----TACGCAAGGCACTCTTAAGAGACTGCGGATGCAAAACGAGAGGAAG	114
QY	1148	GTGGGAGTGAACGTCAAGTCTCTATGAGCCCTTAATGGGTAGGGCTTACACGTAAATACATG	120
Db	1145	GTGGGAGTGAACGTCAAGTCTCTATGAGCCCTTAATGGGTAGGGCTTACACGTAAATACATG	120
QY	1208	GCGGCTAACAGAGGTTTGCCAACCCCGAGGGGAGACTTAATCTCAGAAAACCGGCTGTACT	126
Db	1205	GTCGGTAAACAGAGGCGTGCCAACCCCGAGGTGAGACTTAACCCAGAAAACCGATCTGTACT	126
QY	1268	CCGAGTCCGAGAGTCTCAACTGCACTCCGTAAGTGGGAATCGCTAGTAATCCGCGATCAG	132
Db	1265	CCGAGTCCGAGAGTCTCAACTGCACTCCGTAAGTGGGAATCGCTAGTAATCCGCGATCAG	132
QY	1328	CATGTCCGCGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCACCAATGGGAGTG	138
Db	1325	CATGTCCGCGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCACCAATGGGAGTG	138
QY	1388	GGTTTACACCAAGACAGGTAGTCTAACCGTAAAGAGGGCGCTTGGCCACGCTGAAGAT	1443
Db	1385	GGTTTTCACAGAAAGTAGGTAGCTTAACCGTAAAGAGGGCGCTTAAACCAACGACAGAT	1440

```
/ Publication No. US20030203398A1
/ GENERAL INFORMATION:
/ APPLICANT: Brameucci, Michael
/ TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
/ TITLE OF INVENTION: Industrial Wastewater Bioreactors
/ FILE REFERENCE: BCI033 US NA
/ CURRENT APPLICATION NUMBER: US/10/464,724
/ PRIOR FILING DATE: 2003-06-17
/ PRIOR APPLICATION NUMBER: US/09/735,567
/ PRIOR FILING DATE: 2000-12-13
/ PRIOR APPLICATION NUMBER: 60/171,140
/ PRIOR FILING DATE: 16 DECEMBER 1999
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 4
/ LENGTH: 1453
/ TYPE: DNA
/ ORGANISM: Unknown Organism
/ FEATURE:
/ OTHER INFORMATION: Description of Unknown Organism: Dinitrifying bacteria
/ FEATURE:
/ OTHER INFORMATION: Nucleotide sequence of rDNA corresponding to strain RA2
US-10-464-724-4

Query Match      75.5%; Score 1100; DB 6; Length 1453;
Best Local Similarity 86.1%; Pred. No. 5.5e-296;
Matches 1256; Conservative 0; Mismatches 190; Indels 12; Gaps 3;

QY 1 ATTGAAGCTGGCGCGCATGCTTTACATGCAAGTCCGACGCGACGACGATGCTTCAT 60
DB 1449 ATTGAAGCTGGCGCGCATGCTTTACATGCAAGTCCGACGCGCGCGCA-AC 1394
QY 61 CTGGTGGCGAGTGGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGAT 120
DB 1393 CTGGTGGCGAGCGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGAT 1334
QY 121 ACGCATGGAAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 1333 GCGCGGGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
QY 181 GACCTTGGCGCTTTTGGAGCGCGCGATGATGATGATGATGATGATGATGATGATGAT 240
DB 1273 GCGCTTCAAGCTTTGGAGCGCGCGATGATGATGATGATGATGATGATGATGATGAT 1214
QY 241 CAAGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 1213 CAAGCTTGGCGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1154
QY 301 GCGCGACCTCTTCAAGCGGAGCGAGTGGGAAATTTTGAACATGCGGCAAGCTGATC 360
DB 1153 GCGCGACCTCTTCAAGCGGAGCGAGTGGGAAATTTTGAACATGCGGCAAGCTGATC 1094
QY 361 CAGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 1093 CAGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1034
QY 421 AAAGTTACGGTAAATATGCGAAGCGATGATGATGATGATGATGATGATGATGATGAT 480
DB 1033 AAAGGCTCTCTTATATACAGGCGGATGATGATGATGATGATGATGATGATGATGAT 974
QY 481 TACGTGCGAGAGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 973 TACGTGCGAGAGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914
QY 541 AAAGGTGCGAGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 913 AAAGGTGCGAGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 854
QY 601 TGGCTTTGAATCTAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 853 TGGCTTTGAATCTAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794
QY 661 TGAATGCGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
```

```
DB 793 TGAATGCGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 734
QY 721 GACGCTTATGACGAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 733 GACGCTTATGACGAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 674
QY 781 CTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 673 CTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614
QY 841 AGTTGACCGCGTGGGAGTACGCGTCCGAAAGATTTAACTCAAGAAATTTGACGCGGACCC 900
DB 613 AGTTGACCGCGTGGGAGTACGCGTCCGAAAGATTTAACTCAAGAAATTTGACGCGGACCC 554
QY 901 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 553 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
DB 493 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 434
QY 1018 CTGATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
DB 433 CTGATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 374
QY 373 ACCCTTGCATTAATGCTCATCAATTTGGTGGCACTTTAATGATGATGATGATGATGATGATGAT 1137
DB 1078 ACCCTTGCATTAATGCTCATCAATTTGGTGGCACTTTAATGATGATGATGATGATGATGATGAT 1137
QY 1138 CCGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197
DB 318 CCGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259
QY 1198 TAAATCAATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257
DB 258 TAAATCAATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 199
QY 1258 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
DB 198 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 139
QY 1318 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377
DB 138 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 79
QY 1378 CATGGAGTGGGTTTCAAGAGCAGTATGATGATGATGATGATGATGATGATGATGATGAT 1437
DB 78 CATGGAGTGGGTTTCAAGAGCAGTATGATGATGATGATGATGATGATGATGATGATGAT 19
QY 1438 TGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455
DB 18 CAGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 48
US-10-464-356-4/c
/ Sequence 4, Application US/10464356
/ Publication No. US20030207320A1
/ GENERAL INFORMATION:
/ APPLICANT: Brameucci, Michael
/ TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
/ TITLE OF INVENTION: Industrial Wastewater Bioreactors
/ FILE REFERENCE: BCI033 US NA
/ CURRENT APPLICATION NUMBER: US/10/464,356
/ PRIOR FILING DATE: 2003-06-17
/ PRIOR APPLICATION NUMBER: US/09/735,567
/ PRIOR FILING DATE: 2000-12-13
/ PRIOR APPLICATION NUMBER: 60/171,140
/ PRIOR FILING DATE: 16 DECEMBER 1999
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: Microsoft Office 97
```



```

; SEQ ID NO 4
; LENGTH: 1453
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Dinitrifying bacteria
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of rdna corresponding to strain RA2
US-10-464-356-4
```

```

Query Match          75.5%; Score 1100; DB 6; Length 1453;
Best Local Similarity 86.1%; Pred. No. 5.5e-296;
Matches 1256; Conservative 0; Mismatches 190; Indels 12; Gaps 3;
```

```

QY      1  ATTGAACGCTGCGCGCATGCTTTACACATCAAGTCGAACCGGACGACGCGAGCTTGCAAT 60
DB      1449  ATTGAACGCTGCGCGCATGCTTTACACATCAAGTCGAACCGGACGCGAGCTTGCAAT 1394
QY      61  CTGGTGGCGAGTGGCGGACGCGGTGAGTAATGCAATCGGAACGTATCCAGAAAGAGGGGGT 120
DB      1393  CTGGCGGCGAGCGCGGACGCGGTGAGTAATGCAACATCGGAACGTATCCAGAGTGGGGAT 1334
QY      121  AGCATGCAAAAGATGCTATATACCGCATATCTCTAATGAGGAAAGCAGGGGATCGAA 180
DB      1333  GCGCGGGAAGGCGGATTAATACCGCATATCTCTAATGAGGATGAAAGTGGGGGACCGCA 1274
QY      181  GACCTTGCGCTTTTGGAGCGCGCGATGCTGATTAAGTGAAGTGGGGTAAAGGCTTAC 240
DB      1273  GGCCTACGCGGTTGGAGCGCGCGATGCGATTAAGTGAAGTGGGGTAAAGGCTTAC 1214
QY      241  CAAGGAGAGATGATGATTTGCTGAGAGAGACACGACGACATGGGACTGGGACTGAGAC 300
DB      1213  CAAGCTCGCATTTGATGCTGCTGAGAGATGATGACCACTGGGACTGAGAC 1154
QY      301  GCCCAGACTCTTACGCGGAGCAGCAGTGGGAAATTTTGGCAATGGGCGCAAGCTGATC 360
DB      1153  GCCCAACTCTTACGCGGAGCAGCAGTGGGAAATTTTGGCAATGGGCGCAAGCTGATC 1094
QY      361  CAGCAATGCCGCGTGAAGTGAAGAGCTTTCGGGTTGTAAGCTCTTTCACTGAGAGA 420
DB      1093  CAGCCATTCCGCGTGAAGTGAAGAGCTTTCGGGTTGTAAGCTCTTTGATCGGAACGA 1034
QY      421  AAGGTTCAGCTAATATATGTAACCCATGACCGTATCCAGAGAAGAGACCGGCTAAC 480
DB      1033  AAGGCTCTCTCTAATATACAGGGGCAATATACCGTATCCAGAGAATTAAGACCGGCTAAC 974
QY      481  TACGTCCAGCAGCAGCGCGGTATATACGATGGGTCAAGCGTTAATCGGAATTAATCTGG 540
DB      973  TACGTCCAGCAGCAGCGCGGTATATACGATGGGTCAAGCGTTAATCGGAATTAATCTGG 914
QY      541  AAGGCTGCGCAGCGCGCTTGTATAGTCAGATGTGAATCCCGGCGCTTAACCTTGGGAAT 600
DB      913  AAGGCTGCGCAGCGCGCTTGTATAGTCAGATGTGAATCCCGGCGCTTAACCTTGGGAAT 854
QY      601  TGGGTTGAAACCTACAAAGCTAGAGTGGCAGAGGAGGTGAATTCATGCTGAGCAG 660
DB      853  TGGCTTTGTAAGTGAAGGCTGAGTGGCAGAGGAGGTGAATTCATGCTGAGCAG 794
QY      661  TGAATATGCGTAGATATGGAAGAATCATCGATGGCGAGCAGCTCTTGGGTTAACTACT 720
DB      793  TGAATATGCGTAGATATGCGAGAGAACCATGCGCAAGGCAATCCCTTGGGCTGCACT 734
QY      721  GACGCTCATGACGAAGAGCTGGGGAGCAAAACAGATTAGATACCTTGGTATGTCACGCC 780
DB      733  GACGCTCATGACGAAGAGCTGGGGAGCAAAACAGATTAGATACCTTGGTATGTCACGCC 674
QY      781  CTAAAGATGCTCAATAGTTGTTGGGCTTATTAAGCTTGTGTAACGAAGTACAGGCTGA 840
DB      673  CTAAAGATGCTCAATAGTTGTTGGGAAATTCATTTCTCGTAACGAAGTAAACGCGTGA 614
QY      841  AGTTGACCGCTGGGAGTACGCTCGCAAGATTAAACTCAAGAATTGACGGGAGACC 900
DB      613  AGTTGACCGCTGGGAGTACGCGCCGCAAGTTAAACTCAAGAATTGACGGGAGACC 554
```

```

QY      901  GCACAAGCGGAGGATTAATGATTAATTCAGATGCAAGCGGAAAACTTACTACCTT 960
DB      553  GCACAAGCGGATGATGATGATTTAATTCAGATGCAAGCGGAAAACTTACTACCTTT 494
QY      961  GACATGACGAATTTTCTAGATATGATTAAGTGTCT--TCGGGAACGTAAACAGAGTG 1017
DB      493  GACATGACGAATCTCGAGAGACGAGGATGCTCGAAGAGAGCGCTAACACAGGTG 434
QY      1018  CTGCATGCTGCTGTCAGCTGCTGTCTGAGATGTTGGTTAAGTCCCGCAACGAGCGCA 1077
DB      433  CTGCATGCTGCTGTCAGCTGCTGTCTGAGATGTTGGTTAAGTCCCGCAACGAGCGCA 374
QY      1078  ACCCTTGCAATTAATTCGATATTTGTTGGGCACTTAAAGAGACTGCGGCTGACAA 1137
DB      373  ACCCTTGCAATTAATTCGATATTTGTTGGGCACTTAAAGAGACTGCGGCTGACAA 319
QY      1138  CCGAGGAAGTGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGCTTCAACAG 1197
DB      318  CCGAGGAAGTGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGCTTCAACAG 259
QY      1198  TAATCAATGCGCGCTACAGAGGTTGCCAACCCCGAGGGGAGCTTAATTCAGAAAC 1257
DB      258  TAATCAATGCGCTGTAACAAAGGTTGCCAACCCCGAGGGGAGCTTAATTCAGAAAC 199
QY      1258  GGTGTGATCCGGATCCGAGTCTGCACTGCACTCCGTAAGTGGGAATGCTAGTAAT 1317
DB      198  GGTGTGATCCGGATCCGAGTCTGCACTGCACTCCGTAAGTGGGAATGCTAGTAAT 139
QY      1318  CCGGATCAGCATGTCGCGGTAATACGTTCCGGGTCTTGTACACACCGCGCTACAC 1377
DB      138  CCGGATCAGCATGTCGCGGTAATACGTTCCGGGTCTTGTACACACCGCGCTACAC 79
QY      1378  CATGGAGTGGGTTTCAACGAAGAGATGTTAACCCTAAGAGAGGCGCTTGCCACGG 1437
DB      78  CATGGAGTGGGTTTCAACGAAGAGATGTTAACCCTAAGAGAGGCGCTTGCCACGG 19
QY      1438  TGAGATTCACTGACTGGG 1455
DB      18  CAGGTTCTGTAAGTGGG 1
```

```

RESULT 49
US-10-464-709-4/C
; Sequence 4, Application US/10464709
; Publication No. US20030207321A1
; GENERAL INFORMATION:
; APPLICANT: Bramecci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE REFERENCE: BCI033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,709
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 1453
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Dinitrifying bacteria
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of rdna corresponding to strain RA2
US-10-464-709-4
```

```

Query Match          75.5%; Score 1100; DB 6; Length 1453;
Best Local Similarity 86.1%; Pred. No. 5.5e-296;
Matches 1256; Conservative 0; Mismatches 190; Indels 12; Gaps 3;
```





Db 322 GCCCAGACTCTTAACGGAGGAGCAGCAGTGGGAAATTTGGACAATGGGGGCAACCTGATC 381  
 QY 361 CAGCAATGCGCGGTGAGTAAAGAAAGCCTTGGGTTGTAAGCTCTTTCACTGAGAAAG 420  
 Db 382 CAGCCTGCGCGGTGCTGTAAGAAAGCCTTGGGTTGTAAGCTCTTTCACTGAGAAAG 441  
 QY 421 AAAAGTTAGGTAATTAATCTGTAAGCAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 480  
 Db 442 AATCCCGTGTAAATACCGCGGAGTAAGCAAGGTAAGGTAAGGTAAGGTAAGGTAAG 501  
 QY 481 TACGTCCAGCAGCAGCGGTAAATACGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 540  
 Db 502 TACGTCCAGCAGCAGCGGTAAATACGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 561  
 QY 541 AAAAGTTAGGTAATTAATCTGTAAGCAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 600  
 Db 562 AAAAGTTAGGTAATTAATCTGTAAGCAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 621  
 QY 601 TCGGTTGAAACTCAAAAGCTAGTGTGGCAGAGGAGGTAAGGTAAGGTAAGGTAAGG 660  
 Db 622 GGCATTGAGACTGCAAGACTAGAGTGGCTCAGAGGAGGTAAGGTAAGGTAAGGTAAG 681  
 QY 661 TGAATAGCTAGATGATGAAAGATCGATGAGGTAAGGTAAGGTAAGGTAAGGTAAGG 720  
 Db 682 TGAATAGCTAGATGATGAAAGATCGATGAGGTAAGGTAAGGTAAGGTAAGGTAAGG 741  
 QY 721 GAGCTCATGACGAAAGCGTGGGAGCAAAAGATTAAGTACCTGTAAGTCAAGCC 780  
 Db 742 GAGCTCATGACGAAAGCGTGGGAGCAAAAGATTAAGTACCTGTAAGTCAAGCC 801  
 QY 781 CTAAAGATGTAAGTAAAGTGTGGGCTTAATAGCTGTAAGTAAAGTAAAGTAAAG 840  
 Db 802 CTAAAGATGTAAGTAAAGTGTGGGCTTAATAGCTGTAAGTAAAGTAAAGTAAAG 861  
 QY 841 AGTTAGCCGCTGGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 900  
 Db 862 AATTGACCGCTGGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 921  
 QY 901 GCAAGAGCGGTGATTAATGATTAATGATGCAACGCGAAACCTTAACCTTACCTT 960  
 Db 922 GCAAGAGCGGTGATTAATGATTAATGATGCAACGCGAAACCTTAACCTTACCTT 981  
 QY 961 GACATTAAGCAATTTCTAGAGATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1017  
 Db 982 GACATTAAGCAATTTCTAGAGATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1041  
 QY 1018 CTGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077  
 Db 1042 CTGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1101  
 QY 1078 ACCCTTGTCAATTAATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1137  
 Db 1102 ACCCTTGTCAATTAATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1161  
 QY 1138 CCGGAGAAAGTGGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1197  
 Db 1162 CCGGAGAAAGTGGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1221  
 QY 1198 TAAATCAATGAGCGCTAAGAGGTTGCAACCGCGAGAGGAGTAAAGTAAAGTAAAG 1257  
 Db 1222 TAAATCAATGAGCGCTAAGAGGTTGCAACCGCGAGAGGAGTAAAGTAAAGTAAAG 1281  
 QY 1258 GCGTGTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1317  
 Db 1282 GCGTGTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG 1341  
 QY 1318 CCGGAGTAAAGTGGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1377  
 Db 1342 CCGGAGTAAAGTGGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1401  
 QY 1378 CATGGAGTGGGTTTCAACGAAAGCAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1416  
 Db 1402 CATGGAGTGGGTTTCAACGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1440

RESULT 51  
 US-10-915-740A-8  
 ; Sequence B, Application US/10915740A  
 ; Publication No. US20050191316A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Frazer, Claire M.  
 ; APPLICANT: Hickey, Erin  
 ; APPLICANT: Peterson, Jeremy  
 ; APPLICANT: Tettelein, Herve  
 ; APPLICANT: Venter, J. Craig  
 ; APPLICANT: Massignani, Vega  
 ; APPLICANT: Galeotti, Cesira  
 ; APPLICANT: Mora, Manroza  
 ; APPLICANT: Rateli, Giulio  
 ; APPLICANT: Scarbelli, Maria  
 ; APPLICANT: Scarpato, Vincenzo  
 ; APPLICANT: Rappuoli, Rino  
 ; APPLICANT: Pizza, Mariagrazia  
 ; APPLICANT: Grandi, Guido  
 ; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use  
 ; FILE REFERENCE: 002441.00090  
 ; CURRENT APPLICATION NUMBER: US/10/915,740A  
 ; PRIOR FILING DATE: 2004-08-11  
 ; PRIOR APPLICATION NUMBER: 09/806,866  
 ; PRIOR FILING DATE: 1999-10-08  
 ; PRIOR APPLICATION NUMBER: USSN 60/103,794  
 ; PRIOR FILING DATE: 1998-10-09  
 ; PRIOR APPLICATION NUMBER: USSN 60/132,068  
 ; PRIOR FILING DATE: 1999-04-30  
 ; PRIOR APPLICATION NUMBER: PCT/US99/25373  
 ; PRIOR FILING DATE: 1999-10-08  
 ; NUMBER OF SEQ ID NOS: 1068  
 ; SOFTWARE: Patentin version 3.2  
 ; SEQ ID NO 8  
 ; LENGTH: 20844  
 ; TYPE: DNA  
 ; ORGANISM: Neisseria meningitidis  
 US-10-915-740A-8  
 Query Match 75.2%; Score 1095.8; DB 9; Length 20844;  
 Best Local Similarity 87.0%; Pred. No. 1.5e-294;  
 Matches 1273; Conservative 0; Mismatches 182; Indels 8; Gaps 6;  
 QY 1 ATTTGAACGCTGGCGGATCTTTTACATGCAAGTGAACGGCAGAC--CGATGCTTGC 58  
 Db 18939 ATTGAACGCTGGCGGATCTTTTACATGCAAGTGAACGGCAGACGGCAAGAGCTTGC 18998  
 QY 59 ATCT--GCTGGAGAGTGGGAGCGGGTATGATGATGATGATGATGATGATGATGATG 116  
 Db 18999 TTCTGGGAGTGGGAGTGGGAGCGGGTATGATGATGATGATGATGATGATGATGATG 19058  
 QY 117 GGTACGATCGAAGATGCTTAATACCGATATACCTTAAGAGAGAGAGAGAGAGAG 176  
 Db 19059 GATTAAGTATGAAAGATCAGCTTAATACCGATATACCTTAAGAGAGAGAGAGAGAG 19118  
 QY 177 GAAAGCTTGGGCTTTTGAAGCGGCGATGCTGATTAAGTAAAGTAAAGTAAAGTAAAG 236  
 Db 19119 TTGGGCTTGGGCTTTTGAAGCGGCGATGCTGATTAAGTAAAGTAAAGTAAAGTAAAG 19178  
 QY 237 CTACCAAGCGGAGATCAATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 296  
 Db 19179 CTACCAAGCGGAGATCAATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19238  
 QY 297 CAGGCGCAGACTCTTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 356  
 Db 19239 CAGGCGCAGACTCTTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19297  
 QY 357 GATTCAGCAATGCGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416  
 Db 19298 GATTCAGCAATGCGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19357



DB	Sequence	Position
Db	ACGGCCGCAAGGCTTAAACTC	721
Qy	TGAATTAATTCGATGACGCGA	920
Db	TGGAATTAATTCGATGACGCGA	781
Qy	AGAGATGATTAATTAATTAAT	980
Db	AGAGATTAATTAATTAATTAAT	841
Qy	TCGCTGCTGATGATGATGATG	1037
Db	TCGCTGCTGATGATGATGATG	901
Qy	ATCATTTGTTGGGACCTTTAAT	1097
Db	ATCATTTGTTGGGACCTTTAAT	961
Qy	ACGTCAGTCCCTCAATGCGCC	1157
Db	ACGTCAGTCCCTCAATGCGCC	1021
Qy	GAGGCTTGCCACCCCGGAGGG	1217
Db	GAGGCTTGCCACCCCGGAGGG	1081
Qy	AGCTGCAACTCGACTCCGTGA	1277
Db	AGCTGCAACTCGACTCCGTGA	1141
Qy	GTCGATACGTTCCCGGGCTTT	1337
Db	GTCGATACGTTCCCGGGCTTT	1201
Qy	AGAAAGCAGTATGTTAACCGT	1397
Db	AGAAAGCAGTATGTTAACCGT	1261
Qy	AGAAAGCAGTATGTTAACCGT	1457
Db	AGAAAGCAGTATGTTAACCGT	1321

RESULT 53  
 US-10-464-724-6/c  
 ; Sequence 6, Application US/10464724  
 ; Publication No. US2003020398A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brumack, Michael  
 ; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
 ; TITLE OF INVENTION: Industrial Wastewater Bioreactors  
 ; FILE REFERENCE: ECI033 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/464,724  
 ; PRIOR FILING DATE: 2003-06-17  
 ; PRIOR APPLICATION NUMBER: US/09/735,567  
 ; PRIOR FILING DATE: 2000-12-13  
 ; PRIOR FILING DATE: 16 DECEMBER 1999  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 6  
 ; LENGTH: 1539  
 ; TYPE: DNA  
 ; ORGANISM: Unknown Organism  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Unknown Organism: Brachyomonas  
 ; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RA6  
 US-10-464-724-6

Query Match 74.2%; Score 1080.8; DB 6; Length 1539;  
 Best Local Similarity 85.7%; Pred. No. 1.3e-290;

Matches 1251: Conservative 0; Mismatches 197; Indels 12; Gaps 4			
QY	3	TGAACGCTGGCGCGAGTCGTTTACATCATGCAGAGTCGACGACGACGAGTCGTTGACT	62
Db	1538	TGAACGCTGGCGCGATGCTTACATCATGCAGAGTCGACGACGACGCGCG--TCGTTGCGGAT	1481
QY	63	GGTGGCGCATGTCGCGGACGCGGTGAATTAATGCATCGGAACGTATCCAGAAAGAGGGGGGTAA	122
Db	1480	GGCGCGCATGTCGCGGACGCGGTGAATTAAGCATCGAACGTCCCGGTATGTCCGGGATAGC	1422
QY	123	GCATCGAAAGATGTGCTAATACCCGATATCTTAAAGGAGGAAAGACAGGGATC--GAAA	180
Db	1420	TCGGCGAAAGCCCGATTAATACCCGATAGATCCGTGATGTAAGACAGGGGACCCGCAAG	1361
QY	181	GACCTTGGCTTTTGGAGCGGCGGATGTCGTATTAGCTTAGTTGGTGGGTAAAGGCTTAC	240
Db	1360	GGCTTGGCGCTACTGGAGCGGCGCGATGCAATTAGTAGTTGGTGGGTAAAGGCGCAC	1301
QY	241	CAAGCGCATCATGTAATTGTTGTAAGAGACACACGACCACTGGGACCTGNAACACG	300
Db	1300	CAACCTGCGCATCTGTAGCTGTCTGTAAGAGATATACGCCACATCGGACCTGNAACACG	1241
QY	301	GCCGAGACTCTTAACGGGAGGAGCAGAGTGGGAAATTTTGAACAAATGGGCGCAAGCTGATC	360
Db	1240	GCCGGAACCTCTTAACGGGAGGAGCAGAGTGGGAAATTTTGAACAAATGGGCGCAAGCTGATC	1181
QY	361	CAGCAATGCGCGTGAATGAAGAAAGCCTTCGGGTTGTAAGCTCTTCACTCGAAGAGA	420
Db	1180	CAGCCATGCGCGTGCAGAGATGAAGAGCCTTCGGGTTGTAAGCTCTTTCATCGAAGAGA	1121
QY	421	AAAGTTAAGGTTAATATCGTACCCATGACGGATTCGACAGAAAGAACCCGGCTAAC	480
Db	1120	AAAGGCTCTTTCTAATAAAGGGCACAATGACGGATCCGTAAAGATTAAGCACCGGCTAAC	1061
QY	481	TACGTGCAGAGCGCGGTAAATACGTAGGTGCGAAGCGTTAATCGAATTAATCTGGGCGT	540
Db	1060	TACGTGCAGAGCGCGGTAAATACGTAGGTTGCGAAGCGTTAATCGAATTAATCTGGGCGT	1001
QY	541	AAAGGATGCGGAGCGCGCTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAAT	600
Db	1000	AAAGCGTGCAGGCGCGCTTGTCAAGAACAGTGTGAATCCCGGGCTCAACTTGGGAAC	941
QY	601	TGCGTTTGAATCTAACAAAGCTAAGATGTGGCAGAGGAGGTGGAAATTCATGTGTAGAGAG	660
Db	940	TGCGATTTGTGACTCAAGGCTGAGATACGGGAGGGGGATGTGAATTCGCGTGTAGAGAG	881
QY	661	TGAATGCGTAGATTAATGAAGAACATCGATGGCGAAGGACGCTCTGGGTTTAACT	720
Db	880	TGAATGCGTAGATTAATGGAGGAACACCAATGGCGAAGGACGATCCTCGGCGCTGTACT	821
QY	721	GACGCTCATGCAAGAAACGTTGGGAGCAAAACAGAAATTAGATCCCTGTGTAGTCACGCGC	780
Db	820	GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGAAATTAGATCCCTGTGTAGTCACGCGC	761
QY	781	CTAAACGATGTCAACTAGTTGTTGGGCTTATTAGGCTTGSTTAACGAAGCTTACGCGCTGA	840
Db	760	CTAAACGATGTCAACTGTTGTTGGGGAATTAATTTCTTCACTTAACGAAGCTTACGCGCTGA	701
QY	841	AGTTGACCGCGCTGGGGAGTAGTCCGTCCGCAAGATTTAAACTCAAAAGGAATTTGACGGGACCC	900
Db	700	AGTTGACCGCGCTGGGGAGTAGTCCGCGCGCAAGGTTAAACTCAAAAGGAATTTGACGGGACCC	641
QY	901	GCACAAAGCGGTGATTAATGTGGAATTAATTCGATGCAACGCGAATAAACCCTTACCTACCTT	960
Db	640	GCACAAAGCGGTGATTAATGTGTTTAATTCGATGCAACGCGAATAAACCCTTACCCACCTT	581
QY	961	GACATGTAGCGAATTTTCTAGAGATAGATTAGTGTCT--TGGGAAAGCTTAAACACAGGTG	1011
Db	580	GACATGTATGTAATCCCGCAAGATATGTGGGAGTGTCTGCAAGAGAGGCATTAACAAGGTG	521
QY	1018	CTGATATGGCTGTGTCAGCTGCTGTGTGCGAATGTGGGTTAATGATCCGCAACGAGCGCA	1071
Db	520	CTGATATGGCTGTGTCAGCTGCTGTGTGCGAATGTGGGTTAATGATCCGCAACGAGCGCA	461



|||||  
Db 285 CGGTCTAGTCCGAGTCGAGTCTGCACTCGACTGCGTGAATCGGAATCGTATGAT 226  
|||  
Qy 1318 CGCGATCAGCATGTGCGCGGTGAATAGTTCCGGGTCTTGTACACACCGCCGTCAC 1377  
|||  
Db 225 CGTGATCAGCATGTGCAAGGTGAATAGTTCCGGGTCTTGTACACACCGCCGTCAC 166  
|||  
Qy 1378 CATGGAGTGGGTTTCAACCAAGCAGGTAGTCTTAACCGTAAAGAGGGCGCTTGCACG 1437  
|||  
Db 165 CATGGAGCGGGTCTACCAAGAGAGTAGCTTAACCGCAAGAGAGGGCGCTTGCACG 106  
|||  
Qy 1438 TGAGATTCATGACTGGGGTG 1457  
|||  
Db 105 TGGGGTTCTGACTGGGGTG 86  
|||  
RESULT 55  
US-10-464-709-6/c  
/ Sequence 6, Application US/10464709  
/ Publication No. US20030207321A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Brumucci, Michael  
/ TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
/ TITLE OF INVENTION: Industrial Wastewater Bioreactors  
/ FILE REFERENCE: BC1033 US NA  
/ CURRENT APPLICATION NUMBER: US/10/464,709  
/ PRIOR FILING DATE: 2003-06-17  
/ PRIOR APPLICATION NUMBER: US/09/735,567  
/ PRIOR FILING DATE: 2000-12-13  
/ PRIOR APPLICATION NUMBER: 60/171,140  
/ PRIOR FILING DATE: 16 DECEMBER 1999  
/ NUMBER OF SEQ ID NOS: 30  
/ SOFTWARE: Microsoft Office 97  
/ SEQ ID NO 6  
/ LENGTH: 1539  
/ TYPE: DNA  
/ ORGANISM: Unknown Organism  
/ FEATURE:  
/ OTHER INFORMATION: Description of Unknown Organism: Brachymonas  
/ FEATURE:  
/ OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RA6  
US-10-464-709-6  
Query Match 74.2%; Score 1080.8; DB 6; Length 1539;  
Best Local Similarity 85.7%; Pred. No. 1.3e-290;  
Matches 1251; Conservative 0; Mismatches 197; Indels 12; Gaps 4;  
Qy 3 TGAACCTGGCGGAGCTTTACATGCAAGTCGAACCGGACAGACGATGTTGCATCT 62  
|||  
Db 1538 TGAACCTGGCGGAGCTTTACATGCAAGTCGAACCGGACAGCGC--TCCTTCGGGAT 1481  
|||  
Qy 63 GGTGGCAGAGTGGCGGACGGGTGAGTAATGCATCGGAACGTATCCAGAAAGGGGGTAA 122  
|||  
Db 1480 GGGGGCAGAGTGGCGGACGGGTGAGTAAGATCGGAACGTGCCGGATGAGTGGGATAG 1421  
|||  
Qy 123 GCATGAAAGATGTCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATC--GAAA 180  
|||  
Db 1420 TCGGCAAAAGCCGATTAATACCGCATAGATCCGTGATGAAGACGAGGAGCCCGAAG 1361  
|||  
Qy 181 GACCTTGGCTTTTGAAGCGGCGGATCTGATTAAGTATGTTGGGGTAAAGGGCTTAC 240  
|||  
Db 1360 GGGCTTGGCTTACTGAGCGGCGGATCTGATTAAGTATGTTGGGGTAAAGGGCTTAC 1301  
|||  
Qy 241 CAAGCGACAGATCAGTATGTTGTTCTGAGAGAGACGACACGACACTTGGGACTGAGACAG 300  
|||  
Db 1300 CAAGCTCGGATCTGTAAGTGTCTGAGAGAGATGATAGCCACATCGGAGCTAGACAG 1241  
|||  
Qy 301 GCCCAGACTCTTACGCGAGGACAGCAGTGGGAAATTTTGAACAATGGCGCAGACTGATC 360  
|||  
Db 1240 GCCCAGACTCTTACGCGAGGACAGCAGTGGGAAATTTTGAACAATGGCGCAGACTGATC 1181  
|||  
Qy 361 CAGCATGCCGCTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTTCACTGAGAGAGA 420  
|||

Db 1180 CAGCCATGCCGCGTGCAGAGATGAAGGCTTCGGGTTGTAACTGCTTTTGTACGAAACGA 1121  
|||  
Qy 421 AAAGTTACGGTAATTAATCTGACCCATGAGAGTATGCAAGAAAGACCGGCTAAC 480  
|||  
Db 1120 AAGGCTCTTTTCTAATTAAGAGGACATGACGGTACCGTAAGATTAAGACACCGCTAAC 1061  
|||  
Qy 481 TACGTGCCAGACCGCGGTAATTAAGTGGGTGCAAGCGTTAATGGAATTAAGTGGGCT 540  
|||  
Db 1060 TACGTGCCAGACCGCGGTAATTAAGTGGGTGCAAGCGTTAATGGAATTAAGTGGGCT 1001  
|||  
Qy 541 AAAGGTGCGACGCGCGCTTGTAAATCATGATGTGAATCCCGGAGCTTAACCTGGGAAT 600  
|||  
Db 1000 AAAGGTGCGACGCGCGCTTGTAAATCATGATGTGAATCCCGGAGCTTAACCTGGGAAT 941  
|||  
Qy 601 TCGCTTGAATTAATCAAAAGCTAGAGTGGGCAAGAGAGTGGGAATTCATGTTAGCAG 660  
|||  
Db 940 TCGCATTTGATGATGCAAGGCTGAGATGACGAGAGAGGAGATGAATTCGCTGTAGCAG 881  
|||  
Qy 661 TGAATTCGCTAGATATGTAAGAAATCATGATGGCGAAGGAGGCTCTCTGGTTAACT 720  
|||  
Db 880 TGAATTCGCTAGATATGCGAGAGAAACCGATGGCGAAGGAGTCTCTGGGCTGTACT 821  
|||  
Qy 721 GACGCTCATGACGAAGGCGTGGGAGCAACAGGATTAAGATACCCTGTAGTCCAGCGC 780  
|||  
Db 820 GACGCTCATGACGAAGGCGTGGGAGCAACAGGATTAAGATACCCTGTAGTCCAGCGC 761  
|||  
Qy 781 CTAAACGATGTCACTAGTTGTTGGGCTTATTAAGCTTGGTATGAAGAGCTAAACGCTGA 840  
|||  
Db 760 CTAAACGATGTCACTAGTTGTTGGGATTTATTTCTTCAATGAAGAGCTAAACGCTGA 701  
|||  
Qy 841 AGTTGACGCGCTGGGAGATACGCGTGCAGAAATTAATACTCAAGAGAAATGAAGGAGACC 900  
|||  
Db 700 AGTTGACGCGCTGGGAGATACGCGCGCAAGGTTAAATCTCAAGAGAAATGAAGGAGACC 641  
|||  
Qy 901 GCACAAAGCGGTGATATATGATGATTAATTCATGCAACGCGAATACTTACCTTACCTT 960  
|||  
Db 640 GCACAAAGCGGTGATATGATGATTAATTCATGCAACGCGAATACTTACCTTACCTT 581  
|||  
Qy 961 GACATGTACGGAATTTTCTAGAGATGATTAAGTCT--TCGGGAACGCTAACAGAGTG 1017  
|||  
Db 580 GACATGTATGGAATCCCGGAGAGATGTGGAGTGTCTCGCAAGAGCCATTAACACAGGTG 521  
|||  
Qy 1018 CTGCATGAGCTGTGCTAGAGTCTGATGAGATGTTGGTAAAGTCCCGCAACAGAGCGCA 1077  
|||  
Db 520 CTGCATGAGCTGTGCTAGAGTCTGATGAGATGTTGGTAAAGTCCCGCAACAGAGCGCA 461  
|||  
Qy 1078 ACCCTTGTCAATTAATTCATCAATTAATGATGTTGGGCACTTTAATGAAGCTGCCGTGA 1137  
|||  
Db 460 ACCCTTGTCAATTAATTCATCAATTAATGATGTTGGGCACTTTAATGAGCTGCCGTGA 406  
|||  
Qy 1138 CCGAGAGAAAGTGGGATGATGCTCAAGTCTCAATGGCCCTTAATGGGTAGGGCTTACACAG 1197  
|||  
Db 405 CCGAGAGAAAGTGGGATGATGCTCAAGTCTCAATGGCCCTTAATGGGTAGGGCTTACACAG 346  
|||  
Qy 1198 TAAATCAATGAGGCGGTACAGAGAGGTTGCCAACCCCGAGAGGGAGCTTAATCTCAAGAAC 1257  
|||  
Db 345 TCAATCAATGAGGCGGTACAAAGGACAGCAAGCCCGAGAGTAAAGCAATCCCATTAAGC 286  
|||  
Qy 1258 GCGTCTAGTCCGATCGAGTCTGCACTGACTCCGTGAAGTGGGAATGCTAGTAAT 1317  
|||  
Db 285 GCGTCTAGTCCGATCGAGTCTGCACTGACTCCGTGAAGTGGGAATGCTAGTAAT 226  
|||  
Qy 1318 CGCGATCAGCATGTGCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGTCACAG 1377  
|||  
Db 225 CGTGATCAGCATGTGACCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGTCACAG 166  
|||  
Qy 1378 CATGGAGTGGGTTTCAACCAAGCAGGTAGTCTTAACCGTAAAGAGGGCGCTTGCACAG 1437  
|||  
Db 165 CATGGAGCGGGTCTACCAAGAGAGGTAGCTTAACCGCAAGAGAGGGCGCTTGCACAG 106  
|||  
Qy 1438 TGAGATTCATGACTGGGGTG 1457  
|||  
Db 105 TGGGGTTCTGACTGGGGTG 86  
|||

```
RESULT 56
US-10-464-724-7
; Sequence 7, Application US/10464724
; Publication No. US20030203398A1
; GENERAL INFORMATION:
; APPLICANT: Bramecci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; TITLE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BCI033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,724
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Brachymonas
; OTHER INFORMATION: Nucleotide sequence of the 16S rDNA corresponding to strain RA9
US-10-464-724-7

Query Match      74.1%; Score 1080.2; DB 6; Length 1454;
Best Local Similarity 85.6%; Pred. No. 1.8e-290;
Matches 1251; Conservative 0; Mismatches 198; Indels 12; Gaps 4;

Qy      2 TTGAACCTGCGCGCATCTTATACATGCAAGTCGAACGCGACAGAGATGCTTGATC 61
Db      1 TTGAACCTGCGCGCATCTTATACATGCAAGTCGAACGCGACGCG--TCCTTGGGA 58
Qy      62 TGGTGGGAGTGGGCGGAGCGGGTGAATGATCGATCGAAGCATATCCAGAAAGGGGGGTAA 121
Db      59 TGGCGGAGTGGGCGGAGCGGGTGAATGATCGATCGAAGCATATCCAGAAAGGGGGGTAA 118
Qy      122 CGCATCGAAAGATGTGCTAATACCGCATATATCTTAAAGAGAAAGCAGGCGA--TGGA 179
Db      119 CTGGCGCAAAAGCCGATTTAATACCGCATAGATCCGTGATGAAGCAGGCGACCTCGCA 178
Qy      180 AGACTTGGCTTTTGAAGCGCGCGGATGCTGATTAAGTATGTTGGGGTAAAGGCGTA 239
Db      179 GGGCTTGGCTTCTGAGAGCGCGCGATGTCAATTAGTGTGGGTAAAGGCGCA 238
Qy      240 CCAAGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 299
Db      239 CCAAGGCTGCGATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 298
Qy      300 GGCCGAGACTCTCTACGGGAGGCGACAGTGGGAAATTTTGAACAATGGCGGACCTGAT 359
Db      299 GGCCGAACTCTCTACGGGAGGCGACAGTGGGAAATTTTGAACAATGGCGGACCTGAT 358
Qy      360 CCAGCAATGCGCGCTGACTGAAGAGGCTTGGGTTTAAAGCTCTTTCAGTGAAGAG 419
Db      359 CCAAGCAATGCGCGCTGACTGAAGAGGCTTGGGTTTAAAGCTCTTTCAGTGAAGAG 418
Qy      420 AAAAGGTACGTAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 479
Db      419 AAAAGGCTCTTCTAATAAAGGCGACATGACGTTACGTAAGATTAAGACCGGCTAA 478
Qy      480 CTACGTCCAGAGCGCGGCTAATCTAGAGGTGCAACGCTTAATCGAATTTACTGGCG 539
Db      479 CTACGTCCAGAGCGCGGCTAATCTAGAGGTGCAACGCTTAATCGAATTTACTGGCG 538
Qy      540 TAAAGGTGCGAGCGGCTTGTAAAGTCAAGTGAATCCCGGGCTTAACCTGGGAA 599
Db      539 TAAAGGTGCGAGCGGCTTGTAAAGTCAAGTGAATCCCGGGCTTAACCTGGGAA 598
```

```
Qy      600 TTGCGTTGAACCTAACAAGCTAGATGTGCGAGAGGAGGTGAATTCATGTGTAGA 659
Db      599 CTGCAATTTGATCTGAAGGCTGAGATAGCGAGAGGGAGATGGAATTCGCGGTGAGA 658
Qy      660 GTGAAATGGGTAGATATGAAAGAAATCATGATGCGGAAAGGACCTCTTGGGTAAAC 719
Db      659 GTGAAATGGGTAGATATGCGAGAGAACACCGATGGGAAAGGAGTCCCTGGGCTGTAC 718
Qy      720 TGACGCTATGCAACGAACACCGTGGGAGGAAACAGATTAGATACCTGGTATGCCAGC 779
Db      719 TGACGCTATGCAACGAACACCGTGGGAGGAAACAGATTAGATACCTGGTATGCCAGC 778
Qy      780 CTTAAACAGATGCACTAGTGTGGGCTTATTAAGCTTGTGTAACGAAGCTAACCGGTG 839
Db      779 CTTAAACAGATGCACTAGTGTGGGCTTATTAAGCTTGTGTAACGAAGCTAACCGGTG 838
Qy      840 AAGTTGACCGGCTGGGAGTACCGGTGCAAGATTAAACTCAAGGAATTGACGGGAGC 899
Db      839 AAGTTGACCGGCTGGGAGTACCGGCGCAAGGTTAAACTCAAGGAATTGACGGGAGC 898
Qy      900 CGCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 959
Db      899 CGCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958
Qy      960 TGACATGTAGCGAATTTTCTAGAGATGATTAAGTGTCT--TCGGAAACGCTAACAGGT 1016
Db      959 TGACATGTAGCGAATTTTCTAGAGATGATTAAGTGTCT--TCGGAAACGCTAACAGGT 1018
Qy      1017 GGTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1076
Db      1019 GCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1078
Qy      1077 AACCTTGTGATTAATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1136
Db      1079 AACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1133
Qy      1137 ACCGAGGAAGTGGGAGTGAAGTCAAGTCTGATGAGGCTTATGAGGTGAGGCTTACAC 1196
Db      1134 ACCGAGGAAGTGGGAGTGAAGTCAAGTCTGATGAGGCTTATGAGGTGAGGCTTACAC 1193
Qy      1197 GTAATACATGAGCGGCTGACAGAGGTTGCCAACCCCGAGAGGAGAGTAACTTCAGAAAG 1256
Db      1194 GTCAATACATGAGCGGCTGACAGAGGTTGCCAACCCCGAGAGGAGAGTAACTTCAGAAAG 1253
Qy      1257 CGGCTGTAGTCCGATCGAGTCTGCACTCTCGTGAAGTCCGAAATCGCTAATA 1316
Db      1254 CGGCTGTAGTCCGATCGAGTCTGCACTCTCGTGAAGTCCGAAATCGCTAATA 1313
Qy      1317 TCGCGGATTCAGCATGTGCGGTGAATTCGTTCCCGGCTTGTGAACAACCGCCGCTACA 1376
Db      1314 TCGGTGATTCAGCATGTGCAAGTGAATTCGTTCCCGGCTTGTGAACAACCGCCGCTACA 1373
Qy      1377 CCAAGGAGTGGGTTTCAACAGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 1436
Db      1374 CCAAGGAGTGGGTTTCAACAGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 1433
Qy      1437 GTGAGATTCAATGACTGGGCTG 1457
Db      1434 GTGGGTTTCGTGACTGGGCTG 1454

RESULT 57
US-10-464-356-7
; Sequence 7, Application US/10464356
; Publication No. US20030207320A1
; GENERAL INFORMATION:
; APPLICANT: Bramecci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; TITLE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BCI033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,356
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
```



```

; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Brachymonas
; OTHER INFORMATION: Nucleotide sequence of the 16S rDNA corresponding to strain RA9
US-10-464-356-7

Query Match      74.1%; Score 1080.2; DB 6; Length 1454;
Best Local Similarity 85.6%; Pred. No. 1.8e-290;
Matches 1251; Conservative 0; Mismatches 198; Indels 12; Gaps 4;

QY      2 TTGAAGCGTGGCGGCGATGCTTTACATGCAAGTCCAGCGGACGCGATGCTTGCAATC 61
DB      1 TTGAAGCGTGGCGGCGATGCTTTACATGCAAGTCCAGCGGACGCGATGCTTGCAATC 58
QY      62 TGGTGGCGAGTGGCGGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 121
DB      59 TGGTGGCGAGTGGCGGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 118
QY      122 CGCATGGAAGATGCTGTAATACCGCATATCTCTAAGAGAGAGAGAGAGAGAGAGAGAG 179
DB      119 CTCGGGGAAGCGCGGATTAATACCGCATGATCCCTGATGAGAGAGAGAGAGAGAGAGAG 178
QY      180 AAGCCTTGGCTTTTGAAGCGCGGATGCTGATGATGATGATGATGATGATGATGATGATGAT 239
DB      179 GGGCTTGGCTTTTGAAGCGCGGATGCTGATGATGATGATGATGATGATGATGATGATGAT 238
QY      240 CCAAGCGAGCGATCAGTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 299
DB      239 CCAAGCGTGGATCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 298
QY      300 GGGCGGAGCTCTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
DB      299 GGGCGGAGCTCTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358
QY      360 CCAAGCGATCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419
DB      359 CCAAGCGATCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418
QY      420 AAAAGGTTACGTTAATATCGTACCCATGACGATGATGATGATGATGATGATGATGATGAT 479
DB      419 AAAAGGCTCTTTTAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478
QY      480 CTACGTGCAGAGCGCGGTTAATACGTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
DB      479 CTACGTGCAGAGCGCGGTTAATACGTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 538
QY      540 TAAAGGTTGGCGGCGCTTGTAGTCAAGTGAATATCCCGGGCTTAACTGAGAGAGAGAG 599
DB      539 TAAAGGTTGGCGGCGCTTGTAGTCAAGTGAATATCCCGGGCTTAACTGAGAGAGAGAG 598
QY      600 TTGCGTTTGAATTAACAAGCTAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
DB      599 CTGCACTTGTAGTCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 658
QY      660 GTGAATGCGTAGATATGGAAGAACATGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
DB      659 GTGAATGCGTAGATATGCGAGAGAACATGATGCGGAGAGAGAGAGAGAGAGAGAGAGAG 718
QY      720 TGAAGCTCATGACGAAAGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779
DB      719 TGAAGCTCATGACGAAAGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778
QY      780 CCAAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
DB      780 CCAAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

```

```

DB      779 CCTAAACGATGCAACTGCTGTTGGGATTTATTTCTTCACTAAGAGAGAGAGAGAGAG 838
QY      840 AAGTTACCGCTGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
DB      839 AAGTTACCGCTGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 898
QY      900 CCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
DB      899 CCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
QY      960 TGAACATGAGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016
DB      959 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018
QY      1017 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
DB      1019 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078
QY      1077 AACCTTGTCAATTAATGCTCATCTTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1136
DB      1079 AACCTTGTCAATTAATGCTCATCTTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1133
QY      1137 ACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1196
DB      1134 ACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1193
QY      1197 GTAATGCAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256
DB      1194 GTAATGCAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
QY      1257 CCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1316
DB      1254 CCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
QY      1317 TCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1376
DB      1314 TCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373
QY      1377 CCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1436
DB      1374 CCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1433
QY      1437 GTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
DB      1434 GTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1454

RESULT 58
US-10-464-709-7
; Sequence 7, Application US/10464709
; Publication No. US20030207321A1
; GENERAL INFORMATION:
; APPLICANT: Brumicci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; TITLE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BC1033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,709
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Brachymonas
; OTHER INFORMATION: Nucleotide sequence of the 16S rDNA corresponding to strain RA9
US-10-464-709-7

```



Query Match 74.1%; Score 1080.2; DB 6; Length 1454;  
 Best Local Similarity 85.6%; Pred. No. 1.8e-290;  
 Matches 1251; Conservative 0; Mismatches 198; Indels 12; Gaps 4;

```

QY 2 TTGAACGCTGGCGCATGCTTTTACATTCAGAGTGGACGACCGATCTTGACATC 61
DB 1 TTGAACGCTGGCGCATGCTTTTACATTCAGAGTGGACGACCGATCTTGACATC 58
QY 62 TGGTGGCGAGTGGCGGATGAGTATGATCGGACGTAATCCAGAGAGGAGGATTA 121
DB 59 TGGGCGCGAGTGGCGGATGAGTATGATCGGACGTAATCCAGAGAGGAGGATTA 118
QY 122 CGCATGGAAGAATGCTGTAATACCGCATATATCTTAAGAGGAAACAGAGGGA--TCGA 179
DB 119 CTCGGGAAAGCCGATTAATACCGCATATATCTTGATGATGAACAGAGGATCTGCA 178
QY 180 AGACCTTGGCTTTTGGAGCGCCCATGCTGATTTAGCTAGTTGGTGGGTTAAAGCCCTA 239
DB 179 GGGCTTGGCTTACGAGCGCCCATGCTGATTTAGCTAGTTGGTGGGTTAAAGCCCTA 238
QY 240 CCAAGGCGAGATCGATAGTGGTCTGAGAGGACGACGACCACTGGGACTGAGACAC 299
DB 239 CCAAGCTGCGATCTGATGCTGCTGAGAGGATGATCAAGCCATCGGAGCTGAGACAC 298
QY 300 GGGCCGAGCTCTCTACGAGGAGGAGCAGTGGGAAATTTTGAACATGGGCGCAAGCTGAT 359
DB 299 GGGCCGAACTCTCTACGAGGAGGAGCAGTGGGAAATTTTGAACATGGGCGCAAGCTGAT 358
QY 360 CCAGCAATGCCGCTGAGTGAAGAGGCTTGGGTTTGAAGCTCTTTCAGTCCGAAAG 419
DB 359 CCAGCCATGCCGCTGAGTGAAGAGGCTTGGGTTTGAAGCTCTTTCAGTCCGAAAG 418
QY 420 AAAAGGTTACGTAATTAATCGTGAACCAATGACGTATGACAGAGAAAGCAGGCTTA 479
DB 419 AAAAGGCTCTTTCTAATTAAGAGGCAATGACGTATGACGTATTAAGCAACCGCTTA 478
QY 480 CTACGCGCAGCAGCCGCGTAAATACGTAAGGCTGCAAGCGTTAATCGAATTACTGGGCG 539
DB 479 CTACGCGCAGCAGCCGCGTAAATACGTAAGGCTGCAAGCGTTAATCGAATTACTGGGCG 538
QY 540 TAAAGGCTGCGAGCGGCTTGTAAAGTCAATGTAATCCCGGCTTAACCTGGGAA 599
DB 539 TAAAGGCTGCGAGCGGCTTGTCAAGACAGTGTGAATCCCGGCTTAACCTGGGAA 598
QY 600 TTGCGTTTGAATTAAGAGCTAGAGTGGCGAGGAGGAGTGAATTCATGATGAGCA 659
DB 599 CTGCGATTGACTGCAAGGCTGAGTACGCGAGAGGAGTGAATTCGCGTGTAGCA 658
QY 660 GTGAATGCGTAGATATGAGAGAAACATGATGCGAAGGCGAGCTCTGCGGTTAACAC 719
DB 659 GTGAATGCGTAGATATGCGAGAAACATGCGAAGGCGAGTCTCCCTGGGCTGTAC 718
QY 720 TGAAGCTCATGACGAAAGGCTGGGAGCAACAGATTAAGTACCTGCTAGTCCAGCG 779
DB 719 TGAAGCTCATGACGAAAGGCTGGGAGCAACAGATTAAGTACCTGCTAGTCCAGCG 778
QY 780 CTTAAAGATGTAATGATGTTGTTGGGCTTATAGGCTTGTGAAGCAAGCTAAGCGGTG 839
DB 779 CTTAAAGATGTAATGATGTTGTTGGGATTTATTTCTTGAATGCAAGCTAAGCGGTG 838
QY 840 AAGTTGACCGCTGGGAGTACGGTCCGACAGATTAATACTCAAGAGATTAAGCGGAGCC 899
DB 839 AAGTTGACCGCTGGGAGTACGGGCGCAAGGTTAAACTCAAGAGATTAAGCGGAGCC 898
QY 900 CGACAAAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAATCTTACTTACCTT 959
DB 899 CGACAAAGCGGTGATGATGTTAATTCGATGCAAGCGGAAATCTTACTTACCTT 958
QY 960 TGAAGTACGAAATTTTCTAGAGATGATTAAGTCT---TCGGGAAAGCGTACACAGGT 1016
DB 959 TGAAGTATGAAATCCCGGAGAGATGTGGAGTGTCTCGCAAGAGACCTTAACACAGGT 1018

```

```

QY 1017 GGTGATGCGTGTGCTGACGCTCGTGTGAGATTTGGTTAAGTCCCGCAACGAGCGC 1076
DB 1019 GGTGATGCGTGTGCTGACGCTCGTGTGAGATTTGGTTAAGTCCCGCAACGAGCGC 1078
QY 1077 AACCTTGTCAATTAATTCATCATTTGGGACCTTTAATGAGACTGCGGTGACAA 1136
DB 1079 AACCTTGTCAATTAATTCATCATTTGGGACCTTTAATGAGACTGCGGTGACAA 1133
QY 1137 ACCGAGGAAGTGGGATGACGTACATGCTTCAATGCGCCCTTAATGAGTGGGCTTACAC 1196
DB 1134 ACCGAGGAAGTGGGATGACGTACATGCTTCAATGCGCCCTTAATGAGTGGGCTTACAC 1193
QY 1197 GTAAATCAATGCGGCTGACAGGCTGCAACCGGAGGAGGAGTAACTCAAGAAAG 1256
DB 1194 GTCAATCAATGCGGCTGACAGGCTGCAACCGGAGGAGGAGTAACTCAAGAAAG 1253
QY 1257 GCGCTGTAGTCCGATCGGATCGGATCTGCAACTGACTCCGTGAAGTCCGAATCGTATGA 1316
DB 1254 CCGGTGTAGTCCGATCGGATCGGATCTGCAACTGACTCCGTGAAGTCCGAATCGTATGA 1313
QY 1317 TCGCGATCAGATGCTCGGCTGGAATACGTTCCCGGCTTGTATACACCGCGCCGACAA 1376
DB 1314 TCGGTATCAGATGCTCGGCTGGAATACGTTCCCGGCTTGTATACACCGCGCCGACAA 1373
QY 1377 CGATGGAGTGGGTTTCAACAGACAGTACTTAACCGTAAAGAGGCGCTTGCCACG 1436
DB 1374 CGATGGAGTGGGTTTCAACAGACAGTACTTAACCGTAAAGAGGCGCTTGCCACG 1433
QY 1437 GTGAGATTGATGCTGGGCTG 1457
DB 1434 GTGGGCTTGTGACTGGGCTG 1454

```

# RESULT 59

```

US-10-464-724-2/c
; Sequence 2; Application US/10464724
; Publication No. US20030203398A1
; GENERAL INFORMATION:
; APPLICANT: Brumicci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; TITLE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BC1033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,724
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Brachymonas
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain NBYE13
US-10-464-724-2

```

Query Match 74.1%; Score 1079.6; DB 6; Length 1467;  
 Best Local Similarity 85.6%; Pred. No. 2.7e-290;  
 Matches 1251; Conservative 0; Mismatches 199; Indels 12; Gaps 4;

```

QY 1 ATTGAACGCTGGCGCATGCTTTTACATTCAGAGTGGACGACCGATCTTGACAT 60
DB 1465 ATTGAACGCTGGCGCATGCTTTTACATTCAGAGTGGACGACCGATCTTGACAT 1408
QY 61 CTGGTGGCAGATGGCGGAGCGGTGAGTAAATGCAATCGGAACGTAATCCAGAGGAGGAGTA 120
DB 1407 ATGGCGCGAGTGGCGGAGCGGTGAGTAAATGCAATCGGAACGTAATCCAGAGGAGTA 1348
QY 121 ACGCATCGAAAGATGTGCTTAATCCGCATTAATCTTAAGAGAGAAAGCAGGGA--TCGA 178

```

Db 1347 GTCGCGCGAAGCCGGATTATATCCGCATGAGATCCGTGATGATGAAAGCAGGGGATCTCGA 1288  
Qy 179 AAGACCTTGCGCTTTTGGAGCGCGGAGTCTGATTTAGTACTGTTGGTGGGTAAGGCGCT 238  
Db 1287 AGGGCTTGCGCTACTGAGCGCGCGATGTCAGATTAGGTAGTGGTGGGTAAGGCGCC 1228  
Qy 229 ACCAAGGCGAGCATCATGATGTTGGTCTGAGAGGACGACCGACACACTGGGACTAGACA 298  
Db 1227 ACCAAGCGCTGCATCTGATGCTGCTGAGAGATGATAGCCACATCGGAGCTAGACA 1168  
Qy 299 CGGCGCAGACTCTCTACGCGAGGACAGACAGTGGGAAATTTTGGACAAATGGGCGCAACCTGA 358  
Db 1167 CGGCGCGAAGCTCTCTACGCGAGGACAGACAGTGGGAAATTTTGGACAAATGGGCGCAACCTGA 1108  
Qy 359 TCCAGCAATGCGCGGTGAGTGAAGAGCGCTTCGGGTTGTAAAGCTCTTTCAGTCAAGAA 418  
Db 1107 TCCAGCAATGCGCGGTGAGTGAAGAGCGCTTCGGGTTGTAAAGCTCTTTCAGTCAAGAA 1048  
Qy 419 GAAAAGGTACGCTAAATATGTCATCCCATGACGCTATCGACAGAGAGACCGGCTA 478  
Db 1047 GAAAAGGCTCTTCTTAATTAAGAGGACATGACGCTACCTTAAGATTAAGACCGGCTA 988  
Qy 479 ACTACGTGCGACAGCGCGGTATATACGTAGGGTCAAGCGTTAAATCGGAATTAATGAGG 538  
Db 987 ACTACGTGCGACAGCGCGGTATATACGTAGGGTCAAGCGTTAAATCGGAATTAATGAGG 928  
Qy 539 GTAAAGGTGCGCAGCGCGCTTGTATGTCAGATGTAAATCCCGGCTTAACCTGAGGA 598  
Db 927 GTAAAGCGTGCAGAGCGGCTTGTGCAAGACATGTGTAATCCCGGCTCAACCTGAGGA 868  
Qy 599 ATTGCGTTGAAACTTCAAAAGTATAGTGTGCGACAGAGGAGTGAATTCATGTTGATG 658  
Db 867 ACTGCAATTGTACGTCAAGGCTGAGTACGCGCAGAGGGGATGGAATTCCTCGTGTAGC 808  
Qy 659 AGTGAATGCGTAGAGATATGGAAGAACTCGATGGCGAAGCGACCTCTGGTTAAACA 718  
Db 807 AGTGAATGCGTAGATATGCGAGAGAACCCATGGCGAAGGACGCTCCCTGGGCTGTA 748  
Qy 719 CTGACGCTCATGACGAAAGCGTGGGGAGCAAAACAGATTAGATACCTGTGATGCCAG 778  
Db 747 CTGACGCTCATGACGAAAGCGTGGGGAGCAAAACAGATTAGATACCTGTGATGCCAG 688  
Qy 779 CCTTAAAGATGTCACTAGTGTGGGCTTATAGGCTTGTGAACGAAGTAAAGCGGT 838  
Db 687 CCTTAAAGATGTCACTAGTGTGGGATTTATTTCTTCACTAAGCAAGGCTAACGCT 628  
Qy 839 GAAAGTTGACCGGCTGGGAGTACGGTTCGCAAGATTAAACTCAAAAGAAATTGACGGGAGC 898  
Db 627 GAAAGTTGACCGGCTGGGAGTACGGGCTCGCAAGGTTAAACTCAAAAGAAATTGACGGGAGC 568  
Qy 899 CGGCAAGACGGGTGATATATGATGATTAATTCGATGCAACGCGAAAAACCTTACTTACC 958  
Db 567 CGGCAAGACGGGTGATATGATGATTAATTCGATGCAACGCGAAAAACCTTACTTACC 508  
Qy 959 TTGAACATGAGGAAATTTTCTAGAGATGATTAATGCT---TCGGGAACGCTTAACACAG 1015  
Db 507 TTGAACATGATGAAATCCCGCAGAGATGTGGAGTGTCTGCAAGAGAGCATTAACACAG 448  
Qy 1016 TGCCTCATGCGCTGTGCTGAGCTCGTGTGCTGAGATGTTGGGTTAAAGTCCCGCAAGAGG 1075  
Db 447 TGCCTCATGCGCTGTGCTGAGCTCGTGTGCTGAGATGTTGGGTTAAAGTCCCGCAAGAGG 388  
Qy 1076 CAACCTTGTCATTAATTCGCATCATTTGGTGGGCACTTTAATAGACTGCGCGGTGACA 1135  
Db 387 CAACCTTGTCATCATGTTGCTACGAAA-----GGGCACTCTGATGGGACTGCGCGGTGACA 333  
Qy 1136 AACCGAGAGAGGTGGGATGACGTCAGTCTCATGCGCTTAAATGGTGGGCTTACACA 1195  
Db 332 AACCGAGAGAGGTGGGATGACGTCAGTCTCATGCGCTTAAATGGTGGGCTTACACA 273  
Qy 1195 CGTAAATCAATGGCGCGTACAGAGGTTGCCAACCGCGAGGGGAGCTTAATCTCAGAAA 1255

Db 272 GGTCAATCAATGCGCGGTATCAAAAGGCGACGAGCCCGAGGTGAGGCCAATCCCATMAA 213  
Qy 1256 GCGGCTGTATTCGGGATTCGGAGTCTGCAACTCCGAGAGTGGGATTCGCTAGTA 1315  
Db 212 GCGGCTGTATTCGGGATTCGGAGTCTGCAACTCCGAGAGTGGGATTCGCTAGTA 153  
Qy 1316 ATCCGCGATGACATGTCGCGGTGAATACGTTCCCGGCTTGTATACACACCCCGCTGAC 1375  
Db 152 ATCCGCGATGACATGTCGCGGTGAATACGTTCCCGGCTTGTATACACACCCCGCTGAC 93  
Qy 1376 ACCATGGAGTGGGTTTCCACAGAGAGTATGCTTAAACGTTAAGAGAGGCGCTTGCCAC 1435  
Db 92 ACCATGGAGTGGGTTTCCACAGAGAGTATGCTTAAACGTTAAGAGAGGCGCTTGCCAC 33  
Qy 1436 GGTGAGATTCAATGACTGGGGTG 1457  
Db 32 GGTGAGTTCGTGACTGGGGTG 11

RESULT 60  
US-10-464-356-2/c  
; Sequence 2, Application US/10464356  
; Publication No. US20030207320A1  
; GENERAL INFORMATION:  
; APPLICANT: Brumucci, Michael  
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
; FILE REFERENCE: BCI033 US NA  
; CURRENT APPLICATION NUMBER: US/10/464,356  
; PRIOR FILING DATE: 2003-06-17  
; PRIOR APPLICATION NUMBER: US/09/735,567  
; PRIOR FILING DATE: 2000-12-13  
; PRIOR APPLICATION NUMBER: 60/171,140  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 1467  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Brachymonas  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain NBE13  
US-10-464-356-2

Query Match 74.1%; Score 1079.6; DB 6; Length 1467;  
Best Local Similarity 85.6%; Pred. No. 2.7e-290;  
Matches 1251; Conservative 0; Mismatches 159; Indels 12; Gaps 4;  
Qy 1 ATTTGAACGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGCGATGCTTGCAAT 60  
Db 1465 ATTTGAACGCTGGCGGATGCTTTACATGCAATGCAAGCGGACGCG--TCTTTCGG 1408  
Qy 61 CTGGTGGGAGTGGCGGACGCGGTGATGATCATTCGGAACCTATTCAGAAAGAGGGGGTA 120  
Db 1407 ATGGCGGAGATGGCGGAACGGGTGATGATGATGATGATGATGATGATGATGATGATGAT 1348  
Qy 121 ACGCATCGAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGA--TCGA 178  
Db 1347 GCTCGGCGAAGCCGGATTAATACCGCATGATTCGCGATGAAAGCAGGGGACTCGCA 1288  
Qy 179 AAGACCTTGCGCTTTTGGAGCGCGGATGTGATTTAGTACTGTTGGTGGGTTAAAGCGCT 238  
Db 1287 AGGGCTTGCGCTACTGAGAGCGCGGATGTCAATGATGATGATGATGATGATGATGATGATGAT 1228  
Qy 239 ACCAAGGCGAGCATCATGATGTTGGTCTGAGAGGACGACCGACACACTGGGACTAGACA 298  
Db 1227 ACCAAGCGCTGCATCTGATGCTGCTGAGAGATGATAGCCACATCGGAGCTAGACA 1168  
Qy 299 CGGCGCAGACTCTCTACGCGAGGACAGACAGTGGGAAATTTTGGACAAATGGGCGCAACCTGA 358  
Db 1167 CGGCGCGAAGCTCTCTACGCGAGGACAGACAGTGGGAAATTTTGGACAAATGGGCGCAACCTGA 1108



```
Db      927  GTTAAAGCGTGGCGAGCGGTCTTGCAAGACAGTGCGTAATCCCGGGCTCAACCTGGGA 868
Qy      599  ATTGGCTTTGAACCTACAAAGCTAGAGTGGCAGAGGAGGTGAATTCATGTGTAC 658
Db      867  ACTGCATTGTGACTGCAAGGGTGGAGTACGGCAGAGGGGGAATGAAATCCGGGTGAC 808
Qy      659  AGTGAATCGTAGAGATATGGAAGAACTCCATGGCGAAAGCGCTCTGGTTAA 718
Db      807  AGTGAATCGTAGAGATATGCGAGAAACCGATGCGAAGGAGTCCCTGGGCTGTA 748
Qy      719  CTGAGCGCTATGACGAAAGCGGGGAGCAACAGATTAGTACCTGTGTAGTCAG 778
Db      747  CTGAGCGCTATGACGAAAGCGGGGAGCAACAGATTAGTACCTGTGTAGTCAG 688
Qy      779  CCTTAACGATGCTAGTATTGTTGGCTTATTAGGCTGTGTAAGAACTAACCGCT 838
Db      687  CCTTAACGATGCTAGTATTGTTGGCTTATTAGGCTGTGTAAGAACTAACCGCT 628
Qy      839  GAAATTGACCGCTGGGAGTACGCTGCGAAGTTAAATCTCAAGAAATTGACGGGAC 898
Db      627  GAAATTGACCGCTGGGAGTACGCTGCGAAGTTAAATCTCAAGAAATTGACGGGAC 568
Qy      899  CCGCAACGCGGTGATTTATGTTGATTTATTTGATGCAACGCGAATAACCTTACCC 958
Db      567  CCGCAACGCGGTGATTTATGTTGATTTATTTGATGCAACGCGAATAACCTTACCC 508
Qy      959  TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTCT--TCGGAAAGCTTACACAG 1015
Db      507  TTGACATGTAGTGAATCCCGAGAGATGCGAGTGTGCGAAGAGCCATACACAG 448
Qy      1016  TGTGATGCTGTCTGCTGAGCTGCTGCTGAGATGTTGGTTAAATCCCGCAACAG 1075
Db      447  TGTGATGCTGTCTGCTGAGCTGCTGCTGAGATGTTGGTTAAATCCCGCAACAG 388
Qy      1076  CAACCTTGTCAATTTGTCATCTTTGGTGGGCACTTAAATGAGATCCCGGTGAC 1135
Db      387  CAACCTTGTCAATTTGTCATCTTTGGTGGGCACTTAAATGAGATCCCGGTGAC 333
Qy      1136  AACCGAGGAAAGTGGGAGTGAAGTCAAGTCTCATGCTGCTTATGGTGAAGGCTTAC 1195
Db      332  AACCGAGGAAAGTGGGAGTGAAGTCAAGTCTCATGCTGCTTATGGTGGGCTTAC 273
Qy      1196  CGTAATACATGCGCGGTACAGAGGTTGCCAAACCGCGAGGGGAGCTTATCTCAGAA 1255
Db      272  CGTATACATGCGCGGTACAGAGGTTGCCAAACCGCGAGGGGAGCTTATCTCAGAA 213
Qy      1256  GCGCGTCTAGTCCGATCGGATCGGATCTGCACTCGGTGAAGTCGGAATGCTAGTA 1315
Db      212  GCGCGTCTAGTCCGATCGGATCGGATCTGCACTCGGTGAAGTCGGAATGCTAGTA 153
Qy      1316  ATCGCGATCAGCATGTCGCGGTGAATACGTTCCCGGCTTGTATACACACCGCCGTAC 1375
Db      152  ATCGTGATCAGCATGTCAGCGTGAATACGTTCCCGGCTTGTATACACACCGCCGTAC 93
Qy      1376  ACCATGGAGTGGGTTTACCAAGAGAGGTAGCTTAACTGTAAGAGAGGCGCTTGAC 1435
Db      92  ACCATGGAGGCGGTCTCACACAGAGAGGTAGCTTAACTGTAAGAGAGGCGCTTGAC 33
Qy      1436  GGTGAATTCATGATCGGGGTG 1457
Db      32  GGTGGGTTCTGTGATCGGGGTG 11
```

```
RESULT 62
US-10-168-337A-7
; Sequence 7, Application US/10168337A
; Publication No. US20030170654A1
; GENERAL INFORMATION:
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholtz, Philip; and Blackall,
; TITLE OF INVENTION: Probes and Primers for the Detection of PolypHosphate
; TITLE OF INVENTION: Accumulating Organisms in Wastewater
```

```
; FILE REFERENCE: 002367FC/KF
; CURRENT APPLICATION NUMBER: US/10/168,337A
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1320
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:
; OTHER INFORMATION: PolypHosphate-accumulating organism
US-10-168-337A-7

Query Match      73.9%; Score 1076; DB 6; Length 1320;
Best Local Similarity 89.9%; Pred. No. 2,6e-289;
Matches 1188; Conservative 1; Mismatches 126; Indels 7; Gaps 3;

Qy      1  ATTGAACGCTGGCGCATGCTTTACATGCAAGTCGAAACGGGACAGACGATGCTTGCA 60
Db      1  ATTAAACGCTGGCGCATGCTTTACATGCAAGTCGAAACGGGACAGACGATGCTTGCA 58
Qy      61  CTGTTGGCAGATGGCGGAGCGGGTGAATATGCAATCGGAACGTATCCAGAAAGGGGGTA 120
Db      59  CTGTTGGCAGATGGCGGAGCGGGTGAATATGCAATCGGAACGTATCCAGAAAGGGGGTA 118
Qy      121  AGCGATCGAATAATGCTTAATACCGCATATCTCTAAGAGAGAAAGAGGGGATCGAA 180
Db      119  AGCGAGCGAATAATGCTTAATACCGCATATCTCTAAGAGAGAAAGAGGGGATCGAA 178
Qy      181  GACCTTGCGCTTTGGAGCGCGCATGCTGATTAAGTATGTTGGTGAAGGCTTAC 240
Db      179  GACCTTGCGCTTTGGAGCGCGCATGCTGATTAAGTATGTTGGTGAAGGCTTAC 238
Qy      241  CAAGCGAGCATCAGTAGTTGTTGAGAGAGCAACACCACTGGGACTGAGACAG 300
Db      239  CAAGCGAGCATCAGTAGTTGTTGAGAGAGTATCCGCCACACTGGGACTGAGACAG 298
Qy      301  GCCCAGACTCTTACGGGAGGCGAGTGGGAAATTTGCAATAGGCGGCAAGCCGATC 360
Db      299  GCCCAGACTCTTACGGGAGGCGAGTGGGAAATTTGCAATAGGCGGCAAGCCGATC 358
Qy      361  CAGCAATGCGCGGTGATGAAGAGCCCTTGGGTTGTAAGCTCTTTCAGTCGAGAGA 420
Db      359  CAGCAATGCGCGGTGATGAAGAGCCCTTGGGTTGTAAGCTCTTTCAGTCGAGAGA 418
Qy      421  AAAGTTACGTTAAATATCTGATCCATGACGATATCGACAGAAAGAACCCGCTAAC 480
Db      419  AAATGCTTGGTTAATACCTGATGATGACGTTACCGAATTAAGAACACCGGCTAAC 478
Qy      481  TACGTGCCAGACCGCGGTAAATAGGATGTCAGAGCTTAAATCGGAATTAATCGGCGGT 540
Db      479  TACGTGCCAGACCGCGGTAAATAGGATGTCAGAGCTTAAATCGGAATTAATCGGCGGT 538
Qy      541  AAAGGTGCGGAGGCGGCTTGAAGTCAGATGTGAATCCCGGGCTTAACTGAGGAAT 600
Db      539  AAAGGTGCGGAGGCGGCTTGAAGTCAGATGTGAATCCCGGGCTTAACTGAGGAAT 598
Qy      601  TGGTTTGAATACCAAGCTAGAGTGTGCAAGAGGAGGTGAATTCATGTGTAGCAG 660
Db      599  TGCATTTAGATCGCAAGACTGAGTTTGGCAGAGAGGGGGTGAATTCACAGTGTAGCAG 658
Qy      661  TGAATGCGTAGATATGGAAGAACTGATGTCGAAGGCGAGGCTCTGGGTAACTACT 720
Db      659  TGAATGCGTAGATATGGAAGAACTGATGTCGAAGGCGAGGCTCTGGGTAACTACT 718
Qy      721  GACGCTATGACGAAAGCGGTGGAGCAACAGATTAGATACCTGTAGTCCAGCGC 780
Db      719  GACGCTATGACGAAAGCGGTGGAGCAACAGATTAGATACCTGTAGTCCAGCGC 778
Qy      781  CTAAACGATGTCAACTAGTTGTTGG--GCCTTAATTAAGCTTGGTAAAGAACTAACGCGT 838
Db      779  CTAAACGATGTCAACTAGTTGTTGGAGGGTTAAACCTTTAGTGCCTGATGAACGCGT 838
```



Db 782 CCTAAGCATGTCAGTAGTTGTTGGGATTCATTTCTTACTAAGCTAAGCCGT 841  
Qy 839 GAAGTTGACCGCTGGGAGTACGTCGCAAGTTAACTCAAAGAAATTGACGGGAC 898  
Db 842 GAAGTTGACCGCTGGGAGTACGTCGCAAGTTAACTCAAAGAAATTGACGGGAC 901  
Qy 899 CGGCAAGCGGAGTTATGTCGATTAATTCGATCAACGGGAAACCTTACCTAACCC 958  
Db 902 CGGCAAGCGGAGTACGTCGATTAATTCGATCAACGGGAAACCTTACCTAACCC 961  
Qy 959 TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTCT--TCGGGAAAGCTTAACAG 1015  
Db 962 TTGACATGTAGCGAATTCGCGCTAGAGAGTGGAGTGTGGAAAGAGACCGGGCAG 1021  
Qy 1016 TGTGTCATGCTGTCTGTCAGCTCTGTCGTCTGTCGATGTTGGGTTAAGTCCCGCAAGAG 1075  
Db 1022 TGTGTCATGCTGTCTGTCAGCTCTGTCGTCTGTCGATGTTGGGTTAAGTCCCGCAAGAG 1081  
Qy 1076 CAACCTTGTCTAATTTGTCATCATTTGGTTGGGACCTTATAGACTGCGGGTAC 1135  
Db 1082 CAACCTTGTCTTACTTACTTGC---TACGCAAGAGCATCTTAAGAGACTGCGGGTAC 1136  
Qy 1136 AACCGAGAGAGTGGGAGTACGTCAGTCTCATGCTTATGAGGCTTAC 1195  
Db 1137 AACCGAGAGAGTGGGAGTACGTCAGTCTCATGCTTATGAGGCTTAC 1195  
Qy 1196 CGTAAATACATGCGCGGTACAGAGGTTGCGAACCCGCGAGGGAGCTTATCTCAGAA 1255  
Db 1196 CGTAAATACATGCGCGGTACAGAGGTTGCGAACCCGCGAGGGAGCTTATCTCAGAA 1254  
Qy 1256 GCGCGTCGTAGTCCGATCGGATCTGTCATCTCCGTCGAAATCGGATCGTAACTA 1315  
Db 1255 ACCGATCGTAGTCCGATCGGATCTGTCATCTCCGTCGAAATCGGATCGTAACTA 1314  
Qy 1316 ATCGCGATCAGCATGTCGCGGTGAAATACGTTCCCGGCTTGTACACACCGCCCTAC 1375  
Db 1315 ATCGCGATCAGCATGTCGCGGTGAAATACGTTCCCGGCTTGTACACACCGCCCTAC 1374  
Qy 1376 ACCATGGAGTGGGTTTACCAAGAGAGTACTTACCGTAAAG-AGGGCGCTTGCCA 1434  
Db 1375 ACCATGGAGTGGGTTTACCAAGAGAGTACTTACCGTAAAGAGAGAGAGAGAGAGAG 1434  
Qy 1435 CGGTGATTCATGACTGGG 1454  
Db 1435 CGGTGATTCATGACTGGG 1454

RESULT 64  
US-10-776-767-1  
Sequence 1, Application US/10776767  
Publication No. US20040127687A1  
GENERAL INFORMATION:  
APPLICANT: Casida, Jr., Lester E.  
Falkinham, III, Joseph Oliver  
Cain, Cody C.  
TITLE OF INVENTION: NON-OBLIGATE PREDATORY BACTERIUM  
BURKHOLDERIA CASIDAE AND USBS THEREOF  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1667 K Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/776,767  
FILING DATE: 10-Feb-2004

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/985,846  
FILING DATE: 06-Nov-2001  
APPLICATION NUMBER: 09/063,998  
FILING DATE: 22-Apr-1998  
APPLICATION NUMBER: 60/044,532  
FILING DATE: 23-Apr-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mastrock, Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 8743-006-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)7909090  
TELEFAX: (212)8699741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Genomic DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-776-767-1

Query Match 72.6%; Score 1057.6; DB 7; Length 1495;  
Best Local Similarity 86.4%; Pred. No. 3.7e-284;  
Matches 1262; Conservative 0; Mismatches 184; Indels 14; Gaps 8;

Qy 1 ATTGAACGCTGGCGGATGCTTTTACA-CATGAGATCGAACGGCAGCAGATGCTTGA 59  
Db 3 AATTAATGCTGTTGATGATCTTACGATGATGATGATGATGATGATGATGATGATGAT 62  
Qy 60 TCTGTGCGGAGTGGCGGAGCGGAGTGAATGATGATGATGATGATGATGATGATGATGAT 118  
Db 63 CCGTGTGCGGAGTGGCGGAGCGGAGTGAATGATGATGATGATGATGATGATGATGATGAT 122  
Qy 119 TAAAGCATGAAAGATGCTTAATACCGCATATCTCTTAAGAGAGAGAGAGAGAGATGCA 178  
Db 123 TAGCCCGGAGAAAGCGGATTAATACCGCATATCTCTTAAGAGAGAGAGAGAGAGATGCA 182  
Qy 179 AAGACCTTGGCTTTTGGAGCGCGGATGCTGATGATGATGATGATGATGATGATGATGAT 238  
Db 183 CGGCGCTGCGCTTAATGAGGTTGCGGATGATGATGATGATGATGATGATGATGATGAT 242  
Qy 239 ACCAAGCGACGATCAGTATGTTGTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 298  
Db 243 ACCAAGCGACGATCAGTATGTTGTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301  
Qy 299 CGGCGCAGCTCTTACGAG 358  
Db 302 CGGCGCAGCTCTTACGAG 361  
Qy 359 TCCAGCAATGCGCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418  
Db 362 TCCAGCAATGCGCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421  
Qy 419 GAAAGGTTTACGATTAATATGATGAGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478  
Db 422 GAAATCTTGTGTTTATATAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481  
Qy 479 ACTAGTGCAGAGAGCGCGGTTAATACGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 538  
Db 482 ACTAGTGCAGAGAGCGCGGTTAATACGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541  
Qy 539 GTAAGGAGGAG 598  
Db 542 GTAAGGAGGAG 601  
Qy 599 ATTGCTTTGAAGTAACTAAGAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658  
Db 602 ACTGATTTGTATGAGAGAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661

```
QY 659 AGTGAATGCGTAGAGATATGAAAGAACTATGATGCGGAGGAGCCTCTCGGTTTACA 718
Db 662 AGTGAATGCGTAGAGATATGAAAGAACTATGATGCGGAGGAGCCTCTCGGTTTACA 721
QY 719 CTGACGCTCATGACGAAAGCGTGGGAGGAAACAGATTAAGATACCTGGTAGTCCAGC 778
Db 722 CTGACGCTCATGACGAAAGCGTGGGAGGAAACAGATTAAGATACCTGGTAGTCCAGC 781
QY 779 CCTTAAACGATGTCATGATGTTGTTGGGCTTATTAAGAGCTTGTAAAGAACTAACCGCT 838
Db 782 CCTTAAACGATGTCATGATGTTGTTGGGCTTATTAAGAGCTTGTAAAGAACTAACCGCT 841
QY 839 GAACTTGAACCGCTTGGGAGAGTACCGTGCAGAAATTAATAAATCAAGAAATTTGACGGGAC 898
Db 842 GAACTTGAACCGCTTGGGAGAGTACCGTGCAGAAATTAATAAATCAAGAAATTTGACGGGAC 901
QY 899 CCGGACAAAGCGGTGGATTAATGATTAATTCGATGCAACGGGAAAAAATCTTACCTACCC 958
Db 902 CCGGACAAAGCGGTGGATTAATGATTAATTCGATGCAACGGGAAAAAATCTTACCTACCC 961
QY 959 TTGACATGTAAGCGAATTTTCTAGAGATAGATTAGTCT--TCGGAAACGCTAACACAGG 1015
Db 962 TTGACATGTAAGCGAATTTTCTAGAGATAGATTAGTCT--TCGGAAACGCTAACACAGG 1021
QY 1016 TGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075
Db 1022 TGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081
QY 1076 CAACCTTTGCTATTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1135
Db 1082 CAACCTTTGCTATTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1136
QY 1136 AACCGGAGGAGGTGGGAGTACGATCAAGTCTTATGAGGCTTATGAGGCTTACA 1195
Db 1137 AACCGGAGGAGGTGGGAGTACGATCAAGTCTTATGAGGCTTATGAGGCTTACA 1195
QY 1196 CGTAAATCAATGCGCGTACAGAGGCTTGCACACCGCGGAGGAGCTTAATCTCAAGAA 1255
Db 1196 CGTAAATCAATGCGCGTACAGAGGCTTGCACACCGCGGAGGAGCTTAATCTCAAGAA 1254
QY 1256 GCGGCTGTAAGTCCGAGTCCGAGTCTGCACTCGACTCCGTAAGTCCGTAAGT 1315
Db 1255 ACCATGCTAGTCCGAGTCCGAGTCTGCACTCGACTCCGTAAGTCCGTAAGT 1314
QY 1316 ATGCGGATGCAAGTGTGCGGCTGTAATGTTCCCGGCTTGTACACACCGCGCTAC 1375
Db 1315 ATGCGGATGCAAGTGTGCGGCTGTAATGTTCCCGGCTTGTACACACCGCGCTAC 1374
QY 1376 ACCATGAGAGTGGGTTTACACGAAAGCAGTATCTTAACCGTAAGG-AGGCGCTTGCCA 1434
Db 1375 ACCATGAGAGTGGGTTTACACGAAAGTGGTATCTTAACCGCAAGAAAGAGCGTCCCA 1434
QY 1435 CGGTGAATTCATGACTGGG 1454
Db 1435 CGGTGAATTCATGACTGGG 1454
```

```
RESULT 65
US-09-967-376-1
; Sequence 1, Application US/09967376
; Patent No. US20020106351A1
; GENERAL INFORMATION:
; APPLICANT: WALKER, Harrell L.
; APPLICANT: HIGGINBOTHAM, Lawrence R.
; TITLE OF INVENTION: CONTROL OF CYANOBACTERIA WITH A BACTERIUM
; FILE REFERENCE: 013243-0007
; CURRENT APPLICATION NUMBER: US/09/967,376
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1540
```

```
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Bacterium SG-3; gram negative;
; OTHER INFORMATION: rod-shaped; exhibits flagellate motility; pathogenic to
; OTHER INFORMATION: cyanobacteria and algae; yellow colonies on BG-11 medium suppl.
; OTHER INFORMATION: with tryptic soy
US-09-967-376-1

Query Match 71.9%; Score 1048.2; DB 3; Length 1540;
Best Local Similarity 84.5%; Pred. No. 1,5e-281;
Matches 1239; Conservative 0; Mismatches 218; Indels 10; Gaps 5;

QY 1 ATTGAAGCTGCGGCGCATGCTTTTACATCCAGTGGGAACGGAGACCGGATCTTGAT 60
Db 24 AGTGAAGCTGCGGCGCATGCTTTTACATCCAGTGGGAACGGAGACCGGATCTTGAT 83
QY 61 CT----GTTGCGAGTGGCGGACGCGGTGAGTATGATCGGAACGTAATCCAGAAAGAGGG 116
Db 84 ACTGTGGGTGGCGAGTGGCGGACGCGGTGAGTATGATCGGAACGTAATCCAGAGTGTGGG 143
QY 117 GGTAAAGCTGCGGCGCATGCTTTTACATCCAGTGGGAACGGAGACCGGATCTTGAT 176
Db 144 GATTAAGTAAAGGAACTTATGATTAATCCAGTATACGATCGGAGTAAAGCGGAGATC 203
QY 177 GAAAGACCTTGGCGCTTTTGGAGCGCGGATGCTGATTTAGTACTGTTGGGCTTAAAGC 236
Db 204 GCAAGACCTTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 263
QY 237 CTACCAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 296
Db 264 CAACCAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 323
QY 297 CAGGCGCCAGACTCTTACGAGGAGCAGCAGTGGGAAATTTTGAACAATGGGCGGACCT 356
Db 324 CAGGCGCCAGACTCTTACGAGGAGCAGCAGTGGGAAATTTTGAACAATGGGCGGACCT 383
QY 357 GATTCAGCAATGCCGCTGATGATGAAGAGCCTTGGGCTTGAAGCTCTTTCACTGAG 416
Db 384 GATTCAGCAATGCCGCTGATGATGAAGAGCCTTGGGCTTGAAGCTCTTTGCTCGGA 443
QY 417 AAGAAAGGTTAGGCTTAATATATGATGATGATGATGATGATGATGATGATGATGATGATG 476
Db 444 AAGAAAGGTTAGGCTTAATATATGATGATGATGATGATGATGATGATGATGATGATGATG 503
QY 477 TAACTAGCTCCAGACGCGCGTAAATCCTAGAGGTCAGACGTTAATCGAAATTACG 536
Db 504 TAACTAGCTCCAGACGCGCGTAAATCCTAGAGGTCAGACGTTAATCGAAATTACG 563
QY 537 GCGTAAAGGTTGCGAGCGGCGCTTGTAGTCAAGATGTAATCCCGGCTTAACCTGG 596
Db 564 GCGTAAAGGTTGCGAGCGGCGCTTGTAGTCAAGATGTAATCCCGGCTTAACCTGG 623
QY 597 GAATTTGCTTTGAACCTACAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 656
Db 624 GAATTTGCTTTGAACCTACAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 683
QY 657 GCAGTAAATGCGTAGAGATATGGAAGAAATCATGATGCGGAGGACGCTCTTGGGTTAA 716
Db 684 GCAGTAAATGCGTAGAGATATGGAAGAAATCATGATGCGGAGGACGCTCTTGGGTTAA 743
QY 717 CACTAGGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCA 776
Db 744 CACTAGGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGCA 803
QY 777 CGCCTTAAAGATGATCAATGATTTGGGCTTATTAAG--CTTGTAAAGAAAGTAAAC 834
Db 804 CGCCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 863
QY 835 GCGTGAAGTGAACCGCTGGGAGTACGCTGCGCAAGATTAAACTCAAGAAATTGACGG 894
Db 864 GCGTGAAGTGAACCGCTGGGAGTACGCTGCGCAAGATTAAACTCAAGAAATTGACGG 923
```







```
QY 1132 GACAAACCGGAGAGGTGGGATGACGTCAAGTCTCATGSCCTTATGGGTAGGCTT 1191
    |||
    |||
    |||
Db 1164 GACAAACCGGAGAGGTGGGATGACGTCAAGTCTCATGSCCTTATGGGTAGGCTT 1223
    |||
    |||
    |||
QY 1192 CACACGTATATCAATGGCGGCTACAGAGGTTGCCAACCCGCGAGGGGGGCTTAATCTCA 1251
    |||
    |||
    |||
Db 1224 CACACGTATATCAATGGCGGCTACAGAGGTTGCCAACCCGCGAGGGGGGCTTAATCTCA 1283
    |||
    |||
    |||
QY 1252 GAAAGGCGCTGATAGTCGGATCGGATCTGCAACTGCACTCCGTGAAAGTCGGAATCGCT 1311
    |||
    |||
    |||
Db 1284 GAAACCCCATCTCACTGCTCGGATTTGAGTCTGCAACTGCACTCAATGAAGTCGGAATCGCT 1343
    |||
    |||
    |||
QY 1312 AGTAATCGCGATCAGCA-TGTGCGCGGTGAATAGTTCGCCGGTCTTGTACACACGCCCC 1370
    |||
    |||
    |||
Db 1344 AGTAATCGCGATCAGCA-TGTGCGCGGTGAATAGTTCGCCGGTCTTGTACACACGCCCC 1403
    |||
    |||
    |||
QY 1371 GTCAACACCATGGAGTGGGTTTCAACGAAGCAGGTAGTCTTAACCTTAAGAGAGGGCGCTT 1430
    |||
    |||
    |||
Db 1404 GTCAACACCATGGAGTGGGTTTGTGACACAGAGCAGGTAGTCTTAACCGAAGGGGGCGCTT 1463
    |||
    |||
    |||
QY 1431 GCCACGGTGAGATTCAATGACTGGGGTG 1457
    |||
    |||
    |||
Db 1464 GCCACGGTGAGATTCAATGACTGGGGTG 1490
    |||
    |||
    |||

RESULT 67
US-10-297-465A-1
; Sequence 1, Application US/10297465A
; Publication No. US20040142413A1
; GENERAL INFORMATION:
; APPLICANT: Simpson, Andrew
; APPLICANT: Reinach, Fernando
; APPLICANT: Setubal, Joao
; APPLICANT: Medeiros, Joao
; APPLICANT: Arruda, Paulo
; TITLE OF INVENTION: Isolated Genome of Xylella fastidiosa and Uses Thereof
; FILE REFERENCE: FAPESP 202 US (10213376)
; CURRENT APPLICATION NUMBER: US/10/297,465A
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: PCT/IB01/01618
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,906
; PRIOR FILING DATE: 2001-06-17
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2731748
; TYPE: DNA
; ORGANISM: Xylella fastidiosa
US-10-297-465A-1

Query Match 71.4%; Score 1040.2; DB 7; Length 2731748;
Best Local Similarity 84.1%; Pred. No. 1.6e-278;
Matches 1234; Conservative 0; Mismatches 223; Indels 10; Gaps 5;

QY 1 ATTAAGCGCTGCGCGCATCTTTACATGCAAGTCGAACGCGACGACGATGCTTGAT 60
    |||
    |||
    |||
Db 172301 AGTGAACCTGCGCGCGCATCTTTACATGCAAGTCGAACGCGACGACGATGCTTGAT 172360
    |||
    |||
    |||
QY 61 CT----GGTGGCAGTGGCGGACGCGGTGATGATGATCGGAACCTATCCAGAAAGGGG 116
    |||
    |||
    |||
Db 172361 ACCATGGGTGGCGAGTGGCGGACGCGGTGAGGAATACATCGGAATCTATCGTGGG 172420
    |||
    |||
    |||
QY 117 GGTAAACGATCGAAAGATGTCTAATCCGATATATCTCTAAGAGAGAAAGCAGGGGATC 176
    |||
    |||
    |||
Db 172421 GATTAACGTAGGGAACTTAACGTAAATACCGCATAGACCTACGCGTGAAGAGCAGGGACC 172480
    |||
    |||
    |||
QY 177 GAANAACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTAAAGGC 236
    |||
    |||
    |||
Db 172481 TTAGGGCCCTTGTGCGATTGAGTGAAGCCAGATGTCGATTAGCTTAGTTGGTGAAGGC 172540
    |||
    |||
    |||
QY 237 CTACCAAGCGCAGATCAGTAGTGTGTGAAGAGACGACCAAGCACACTGGGACTGAGA 296
    |||
    |||
    |||
```

```
Db 172541 TCACCAAGCGCAGATCAGTAGTGTGTGAAGAGATGATCAAGCACACTGGAACCTGAGA 172600
    |||
    |||
    |||
QY 297 CACGCGCCAGACTCTCTACGGAGAGCAGCAGTGGGGAATTTTGGACAAATGGGGCGAACGCT 356
    |||
    |||
    |||
Db 172601 CACGCGCTCAGACTCTCTACGGAGAGCAGCAGTGGGGAATTTTGGACAAATGGGGCGAACGCT 172660
    |||
    |||
    |||
QY 357 GATCCAGCAATGCCCGGTGAGTGAAGAGGCCCTTGGGTTGTAAAGCTCTTTCAGTCGAG 416
    |||
    |||
    |||
Db 172661 GATCCAGCAATGCCCGGTGAGTGAAGAGGCCCTTGGGTTGTAAAGCTCTTTCAGTCGAG 172720
    |||
    |||
    |||
QY 417 AAGAAAGGTTACGTTAAATTAATCTGACCAATGACGATATGACAGAAAGAACACCGGC 476
    |||
    |||
    |||
Db 172721 AAGAAAGGTTACGTTAAATTAATCTGACCAATGACGATATGACAGAAAGAACACCGGC 172780
    |||
    |||
    |||
QY 477 TAACTACGTGCGACAGCGCGGTAAATACGTAGGGTGCAGAGCTTAATGGAAATTAATCTG 536
    |||
    |||
    |||
Db 172781 TAACTACGTGCGACAGCGCGGTAAATACGTAGGGTGCAGAGCTTAATGGAAATTAATCTG 172840
    |||
    |||
    |||
QY 537 GCGTAAAGGTTGGCGAGCGCGCTTGTAAAGTCAAGTGAATCCCGGGCTTAATCTG 596
    |||
    |||
    |||
Db 172841 GCGTAAAGGTTGGCGAGCGCGCTTGTAAAGTCAAGTGAATCCCGGGCTTAATCTG 172900
    |||
    |||
    |||
QY 597 GAATGCGTTTGAATTAACAAAGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGA 656
    |||
    |||
    |||
Db 172901 GAATGCGAGTGAATTAACAAAGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGA 172960
    |||
    |||
    |||
QY 657 GCAAGTAATGCTAGAGATATGAAGAAATCATGATGCGAGCGACGCTCTGGGTAA 716
    |||
    |||
    |||
Db 172961 GCAAGTAATGCTAGAGATATGAAGAAATCATGATGCGAGCGACGCTCTGGGTAA 173020
    |||
    |||
    |||
QY 717 CACTGACCTCATGCAAGAAAGCTGGGAGCAAAACAGATTGATACCTGTGATGCCA 776
    |||
    |||
    |||
Db 173021 CACTGACCTCATGCAAGAAAGCTGGGAGCAAAACAGATTGATACCTGTGATGCCA 173080
    |||
    |||
    |||
QY 777 CGCCCTAAAGATGTCAACTAGTTGTTGGCTTATTAAG--CTTGGTAAAGAACTTAC 834
    |||
    |||
    |||
Db 173081 CGCCCTAAAGATGTCAACTAGTTGTTGGCTTATTAAG--CTTGGTAAAGAACTTAC 173140
    |||
    |||
    |||
QY 835 GCGTGAAGTTGACCGCTGGGAGATACGTCGCAAGTTAAACTCAAGAAATTTGACG 894
    |||
    |||
    |||
Db 173141 GCGTGAAGTTGACCGCTGGGAGATACGTCGCAAGTTAAACTCAAGAAATTTGACG 173200
    |||
    |||
    |||
QY 895 GAAACCGCACAAGCGGTGATTAATGTGATTAATTTGATGCCAACGCGAAACCTTACCT 954
    |||
    |||
    |||
Db 173201 GAAACCGCACAAGCGGTGATTAATGTGATTAATTTGATGCCAACGCGAAACCTTACCT 173260
    |||
    |||
    |||
QY 955 ACCCTTGACATGTAAGCAATTTCTAGAGATGATTAAGT--CTTGGGAAACGCTAACACA 1013
    |||
    |||
    |||
Db 173261 ACCCTTGACATGTAAGCAATTTCTAGAGATGATTAAGT--CTTGGGAAACGCTAACACA 173320
    |||
    |||
    |||
QY 1014 GGTGCTGATGAGTGTGTCAGCTCGTGTGTCGATGATGTTGGTTAAATCCCGCAACGAG 1073
    |||
    |||
    |||
Db 173321 GGTGCTGATGAGTGTGTCAGCTCGTGTGTCGATGATGTTGGTTAAATCCCGCAACGAG 173380
    |||
    |||
    |||
QY 1074 CGCAACCTTGTCAATTAATGTCATCA--TTTGGTTGGGCACTTAATGACCTGCGGCT 1131
    |||
    |||
    |||
Db 173381 CGCAACCTTGTCAATTAATGTCATCA--TTTGGTTGGGCACTTAATGACCTGCGGCT 173440
    |||
    |||
    |||
QY 1132 GACAAACCGGAGAGAGTGGGGAATGAGTCAAGTCCATGAGCCCTTATGGGTAGGGCTT 1191
    |||
    |||
    |||
Db 173441 GACAAACCGGAGAGAGTGGGGAATGAGTCAAGTCCATGAGCCCTTATGGGTAGGGCTT 173500
    |||
    |||
    |||
QY 1192 CACACGTATATCAATGGCGGCTACAGAGGTTGCCAACCCGCGAGGGGGGCTTAATCTCA 1251
    |||
    |||
    |||
Db 173501 CACACGTATATCAATGGCGGCTACAGAGGTTGCCAACCCGCGAGGGGGGCTTAATCTCA 173560
    |||
    |||
    |||
QY 1252 GAAAGCGCTGATAGTCGGATCGGATCTGCAACTGCACTCCGTGAAAGTCGGAATCGCT 1311
    |||
    |||
    |||
Db 173561 GAAACCCCATCTCACTGCTCGGATTTGAGTCTGCAACTGCACTCAATGAAGTTGGAATGCT 173620
    |||
    |||
    |||
QY 1312 AGTAATGCGGATGAGCA-TGTGCGCGGTGAATAGTTCGCCGGTCTTGTACACACGCCCC 1370
    |||
    |||
    |||
```

Db 173621 AGTAATCGCAGATCAGCATTCGTGCGGTGAATACGTTCCCGGCGCTTGTACACACCGCCC 173680  
QY 1371 GTACACCATGGAGAGTGGCTTTACACGAGCAGTATGTCTTAACCGTAAGAGGGGCGCTT 1430  
Db 173681 GTACACCATGGAGAGTGGCTTTGACACGAGCAGTATGTCTTAACCGGAGAGGGGCGCTT 173740  
QY 1431 GCCACGGTGAATTCATGACTGGGGTGG 1457  
Db 173741 GCCACGGTGTGGCGCATGACTGGGGTGG 173767

## RESULT 68

US-10-418-861B-53  
; Sequence 53, Application US/1041861B  
; Publication No. US20040010131A1  
; GENERAL INFORMATION:  
; APPLICANT: da Silva, Ana Claudia Raseira  
; APPLICANT: Farah, Shaker Chuck  
; APPLICANT: Quaggio, Ronaldo Bento  
; APPLICANT: Reinach, Fernando de Castro  
; APPLICANT: Ferro, Jesus Aparecido  
; APPLICANT: De Oliveira, Julio Cezar Franco  
; APPLICANT: Secubal Joao C.  
; APPLICANT: Furlan, Luiz Roberto  
; TITLE OF INVENTION: Isolated xanthomonas nucleic acid molecules, proteins encoded the  
; TITLE OF INVENTION: uses thereof  
; FILE REFERENCE: FAPESP 205.1 US  
; CURRENT APPLICATION NUMBER: US/10/418, 861B  
; CURRENT FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: US 60/374, 620  
; PRIOR FILING DATE: 2002-04-22  
; NUMBER OF SEQ ID NOS: 85  
; SEQ ID NO 53  
; LENGTH: 1547  
; TYPE: DNA  
; ORGANISM: Xanthomonas  
; FEATURE:  
US-10-418-861B-53

Query Match 70.7%; Score 1030.2; DB 6; Length 1547;  
Best Local Similarity 83.7%; Pred. No. 1.6e-276;  
Matches 122; Conservative 0; Mismatches 228; Indels 12; Gaps 5;

QY 1 ATTGAACGCTGGGCGCATGCTTTACATGCAAGTCAAGCGCAGCAGCGATGCTTGCAT 60  
Db 28 AGTGAACGCTGGGCGCATGCTTTACATGCAAGTCAAGCGCAGCAGCGATGCTTGCAT 87  
QY 61 CT-----GGTGGCGAATGGCGGAGCGGGTGAATGATCGGAACGATCCAGAAAGAG 114  
Db 88 CTCTTATGGGTGGCGAGTGGCGGAGCGGGTGAATGATCGGAATCTACTCTTTGCTGG 147  
QY 115 GGGGTAACGCGATCGAAAGATGTGTATATACCGCATATATCTTAAGAGAAACAGAGGGA 174  
Db 148 GGGATTAACGTAAGGAAACCTTACGCTAATACCGCATTAACGCGTGAAGAGCGAGGA 207  
QY 175 TCGAAGAAGCTTGGCGCTTTTGGAGCGCGCGATGCTGATTAAGCTAAGTTGGTGGGTAAG 234  
Db 208 CTTTCGGGCTTCCGCGGCTTGAATGAGCCGATGCTGATTAAGCTAAGTTGGCGGGGTAAG 267  
QY 235 GCGTACCAAGGCGCATGCTAGTATGTTGTTGAGAGAGACGACCGACCACTGGAGCTGA 294  
Db 268 GCCCACCMAAGGCGACATCGGTAGCTGTCTGAGAGGATGATCAGCCACTGGAACCTGA 327  
QY 295 GACACGGCCGAGACTCTTAAGGAGGAGCAGAGTGGGAAATTTTGAACAATGGCGCAGC 354  
Db 328 GACACGGTCCAGACTCTTAAGGAGGAGCAGAGTGGGAAATTTTGAACAATGGCGCAGC 387  
QY 355 CTGATCCAGCAATGCGCGGTGAGTGAAGAAAGCCCTTCGSGTTGAAAGCTTTCAAGTCG 414  
Db 388 CTGATCCAGCAATGCGCGGTGAGTGAAGAAAGCCCTTCGSGTTGAAAGCTTTTGTGG 447  
QY 415 AGAAGAAAGGTTACGTTAATATGCTGACCCATGACGATTCGACAGAAAGACACCG 474

Db 448 GAAAGAAAGCAGTGGTTAATACCGCATTTGTTCTGACCGGTCCCAAAAGAAATAGACACCG 507  
QY 475 GCTAATAGTCCGACGACGCGCGGTAATACGTAAGGTTGCAAGCGTTATCGGAATTACT 534  
Db 508 GCTAATAGTCCGACGACGCGCGGTAATACGTAAGGTTGCAAGCGTTATCGGAATTACT 567  
QY 535 GGGCGTAAAGGTTGCGCAGCGCGCTTGAATGATGATGTAATCCCGGCGTTAACT 594  
Db 568 GGGCGTAAAGGTTGCGCAGCGCGCTTGAATGATGATGTAATCCCGGCGTTAACT 627  
QY 595 GGGATTTGCGTTGAAACTTAACAGCTGATGTTGCGCAGAGGAGTGGAAATTTCAATGG 654  
Db 628 GGGAAATGAGTGAATGATGTTGCGTCACTGAGTGTGTGAGAGGTTAGCGGAATTTCCCGGTG 687  
QY 655 TAGCATGAAATGCGTGAATGATGTAAGAAACATGATGCGCAAGCAGCGCTTCGGGTT 714  
Db 688 TAGCATGAAATGCGTGAATGATGTAAGAAACATGATGCGCAAGCAGCGCTTCGGGTT 747  
QY 715 AACACTGACGTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAAGTACCTGGTATGTC 774  
Db 748 AACACTGACGTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAAGTACCTGGTATGTC 807  
QY 775 CAGCGCTTAACGATGTCATGATGTTGTTGGCTTATTAAGCTTG--GTAAAGAGCTTA 832  
Db 808 CAGCGCTTAACGATGTCATGATGTTGTTGGCTTATTAAGCTTG--GTAAAGAGCTTA 867  
QY 833 AGCGTGAAGTGAACCGCTGGGAGTACGCTGCGCAATTTAACTCAAGAAATGAC 892  
Db 868 AGCGTGAAGTGAACCGCTGGGAGTACGCTGCGCAATTTAACTCAAGAAATGAC 927  
QY 893 GGGGACCGGACAGAGGCGGTGATTAATGATGTTAATGATGCAACGCAAAACCTTAC 952  
Db 928 GGGGACCGGACAGAGGCGGTGATTAATGATGTTAATGATGCAACGCAAAACCTTAC 987  
QY 953 CTACCTTGACATGTAAGCAATTTTCTAGAGTAATGATG--CTTCGGAACGCTTACA 1011  
Db 988 CTGCTTGACATGTAAGCAATTTTCTAGAGTAATGATG--CTTCGGAACGCTTACA 1047  
QY 1012 CAGGTGCTGATGCTGCTGTGTAAGTCCGCTGTAAGTGGTGAATGTCGCCAAGC 1071  
Db 1048 CAGGTGCTGATGCTGCTGTGTAAGTCCGCTGTAAGTGGTGAATGTCGCCAAGC 1107  
QY 1072 AGGCAACCTTGTATTAATGATGCACTA--TTTGGTGGGCACTTTAATGAGACTGGCG 1129  
Db 1108 AGGCAACCTTGTATTAATGATGCACTA--TTTGGTGGGCACTTTAATGAGACTGGCG 1167  
QY 1130 GTGACAAACCGAGAGAGGTGGGATGACGTCAAGTCTCATGAGCCTTATGGGTAAGGC 1189  
Db 1168 GTGACAAACCGAGAGAGGTGGGATGACGTCAAGTCTCATGAGCCTTATGGGTAAGGC 1227  
QY 1190 TTCAACGTAATTAATGAGTGGCGGTACAGAGGTTGCAACCGCGAGGGGAGCTAATCT 1249  
Db 1228 TACACGCTAATTAATGAGTGGGAGCAGAGGCTGCAACCGCGAGGGGAGCTAATCT 1287  
QY 1250 CAGAAAGCGCGTGAATGCGGATGCGGATGCGAATCTGCACTCGGTGAAGTCGGAATGC 1309  
Db 1288 CAGAAAGCGCGTGAATGCGGATGCGGATGCGAATCTGCACTCGGTGAAGTCGGAATGC 1347  
QY 1310 CTAGTAAATGCGGATGACGA--TGTGCGGTTGAATAGTTCCCGGCTTTGTACACACCGC 1368  
Db 1348 CTAGTAAATGCGGATGACGA--TGTGCGGTTGAATAGTTCCCGGCTTTGTACACACCGC 1407  
QY 1369 CCGTCAACCATGGAAGTGGGTTTCAACGAAGCAGTATGTTAAACCGTAAGAGGGCGC 1428  
Db 1408 CCGTCAACCATGGAAGTGGGTTTCAACGAAGCAGTATGTTAAACCGTAAGAGGGCGC 1467  
QY 1429 TTGCAACGTTGATTAATGACTGGGGTGG 1457  
Db 1468 TTGCAACGTTGATTAATGACTGGGGTGG 1496

## RESULT 69

```
US-10-464-724-3
; Sequence 3, Application US/10464724
; Publication No. US20030203398A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE REFERENCE: BCI033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,724
; CURRENT FILING DATE: 2003-06-17
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Denitrifying bacteria
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RA1
US-10-464-724-3
Query Match 70.2%; Score 1023.2; DB 6; Length 1450;
Match Local Similarity 84.3%; Pred. No. 1.4e-274;
Matches 1215; Conservative 0; Mismatches 213; Indels 14; Gaps 5;

QY 1 ATTGAAGCTGGCGCGCATGCTTTACACATGCAAGTCGAACGCGACGACGAGTCTTGAT 60
DB 3 ATTGAAGCTGGCGCGCATGCTTTACACATGCAAGTCGAACGCGACGCGCGGCGCA---AC 58
QY 61 CTGGTGGCGAGTGGCGGACGCGGTGATGATGATCGGAACTGATCCGAAAGAGGGGGGTA 120
DB 59 CTGGCGCGCGAGCGCGGACGCGGTGATGATGATCGGAACTGATCCGAAAGAGGGGGGTA 118
QY 121 ACGGATGGAAGATGTGCTAATACCGCATACTCTAAGAGAAAGCAGGGGATCGGAA 180
DB 119 GCCCGCGAAGACCGCGATTATACCGCATGTGATCTGAGATGAAAGTGGGGGACCCGAA 178
QY 181 GACCTTCGCTTTTGAAGCGCGCCGATGTCTGATTAGCTAGTGGTGGGATTAAGGCTTAC 240
DB 179 GCGCTCAAGCGCTTTGAGCGCGCCGATGGCAATTAGGATTTGGTGGGGTAAAGGCTTAC 238
QY 241 CAAGCGCAGATCAGTATGTTGCTGAGAGACGACACGACCACTGGGATCTGAGACACG 300
DB 239 CAAGCTGCGATCTGTAGCTGTCTGAGAGGATGATCAGCCACTGGGATCTGAGACACG 298
QY 301 GCCCAGATCTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 299 GCCCAGATCTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 358
QY 361 CAGCAATGCGCGTGAAGTGAAGAGGCGCTTGGGTTGTAAGCTCTTTCAGTCCGAGAGA 420
DB 359 CAGCCATTCCCGGTGACGAGAGGAGGCGCTTGGGTTGTAAGCTCTTTCAGTCCGAGAGA 418
QY 421 AAAGGTTACGTTAATATATCGTGAACCATGACGCTATGACAGAGAGAGAGACCGGCTAAC 480
DB 419 AAAGGCTCTCTAATATACAGGGGGCATATGACGCTATGAGAAATAGACACCGGCTAAC 478
QY 481 TAGCTGCGACGACCGCGGTGATATAGTATGCTGCAAGGCTTAATCGGAAATTAATCGGGGT 540
DB 479 TAGCTGCGACGACCGCGGTGATATAGTATGCTGCAAGGCTTAATCGGAAATTAATCGGGGT 538
QY 541 AAAGGTTACGTTAATATATCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 600
DB 539 AAAGGTTACGTTAATATATCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 598
QY 601 TCGCTTTGAACTACAAAGCTAGAGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 599 TCGCTTTGAACTACAAAGCTAGAGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 658
```

```
QY 661 TGAATGCTAGATATGAGAGAACATCATGTGCGAAGGACGCTCTGGTTAACA 720
DB 659 TGAATGCTAGATATGCGAGAGAACACCGATGTGTGAAGGCATCCCGTGGGATCATCT 718
QY 721 GACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
DB 719 GACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 778
QY 781 CT-AAACGATGCACTAGTTGTTGGGCTTATTAAGCTTGGTATGAGAGCTTAAGCGCGT 839
DB 779 CTAAACGATGCACTAGTTGTTGGGCTTATTAAGCTTGGTATGAGAGCTTAAGCGCGT 838
QY 840 AAGTTACCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 899
DB 839 AAGTTACCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 898
QY 900 CGCAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
DB 899 CGCAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
QY 960 TGACATGATGCGAATTTTCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016
DB 959 TGACATGATGCGAATTTTCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018
QY 1017 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1076
DB 1019 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1078
QY 1077 AACCTTGATTAATTTGATCATTTGTTGGGCACTTAAATGAGATGCTCCGATGACAA 1136
DB 1079 AACCTTGATTAATTTGATCATTTGTTGGGCACTTAAATGAGATGCTCCGATGACAA 1133
QY 1137 ACCGAGAGAGTGGGAGTACGTCATGCTTCAATGAGCCCTTAATGAGTGGTACGCTTACAC 1196
DB 1134 ACCGAGAGAGTGGGAGTACGTCATGCTTCAATGAGCCCTTAATGAGTGGTACGCTTACAC 1193
QY 1197 GTAATACATGCGCGCTACAGAGGTTGCCAACCCCGAGAGGAGGAGGAGGAGGAGGAG 1256
DB 1194 GTAATACATGCGCGCTACAGAGGTTGCCAACCCCGAGAGGAGGAGGAGGAGGAGGAG 1253
QY 1257 CGGCTGTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCC 1316
DB 1254 CGGCTGTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCC 1313
QY 1317 TCGCGATCAGCATGTGCGCGGTGAATACGTT-CCCGGCTTGTACACACCGCCGCTAC 1375
DB 1314 TCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGCTTGTGAACACCGCCGCTTCA 1373
QY 1376 ACCATGGAGTGGGTTTACCAAGAGGATGATCTAACGTTAAGAGGCGCTTGGCAC 1435
DB 1374 ACACATGAGAGCGGTTTCCCAAGAGGATGATCTAACCGCAAGAGGCGGATTAACAC 1433
QY 1436 GG 1437
DB 1434 GG 1435

RESULT 70
US-10-464-356-3
; Sequence 3, Application US/10464356
; Publication No. US20030207320A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE REFERENCE: BCI033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,356
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
```

```
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Denitrifying bacteria
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RAI
; US-10-464-356-3
```

Query Match 70.2%; Score 1021.2; DB 6; Length 1450;

Best Local Similarity 84.3%; Pred. No. 1.4e-274; Mismatches 213; Indels 14; Gaps 5;

Matches 1215; Conservative 0;

```
1 ATTGAACGCTGCGGCGGCGGCTTTTACACATGCAAGTCGACGCGGAGCAGCGGATCTTGCAAT 60
3 ATTGAACGCTGCGGCGGCGGCTTTTACACATGCAAGTCGACGCGGAGCAGCGGATCTTGCAAT 58
61 CTGGTGGCGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
59 CTGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 118
121 ACCGATCGAAAGATGCTGCTAATACCGCATATCTTCTAGAGGAGGAAAGCAGGCGGATCGAAA 180
119 GCGCGGCGGAGAACCGGATTAATACCGCATGCTGATGAGAGTGAAGTGGGCGGAGCCGCA 178
181 GACCTTGGCGCTTTTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
179 GCGCTCAACGCGGCTTTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 238
241 CAAGGCGAGCATCATGATGTTGCTGAGAGAGAGCAGCAGCAGCATCTGGGATCTGAGACAG 300
239 CAAGCGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 298
301 GCGCGAGCTCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 360
299 GCGCGAGCTCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 358
361 CAGCAATGCGCGCTGAGTGAAGAGGCGCTTCCGGTGTGAAGCTCTTTCAGTGAAGAGA 420
359 CAGCCATTCCGCGTGGAGAGAGAGGCGCTTCCGGTGTGAAGCTCTTTCAGTGAAGAGA 418
421 AAGGTTACGCTAATATATGTAATCCATGACGCTATCCAGAGAGAGAGAGAGAGAGAGAGAG 480
419 AAGGCTCTCTCTAATACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 478
481 TAGCTGCGAGCAGCGCGGCTAATACGTAAGGCTGCAAGCGTAAATCCGGAATTAATCTGGGCG 540
479 TAGCTGCGAGCAGCGCGGCTAATACGTAAGGCTGCAAGCGTAAATCCGGAATTAATCTGGGCG 538
541 AAGGCTGCGAGCGGCGCTTGTAAATGTCAGATGTAATCCCGGCGTAACTTGAGGAT 600
539 AAGCGTGGCGAGCGGCTTGTAAATGTCAGATGTAATCCCGGCGTAACTTGAGGAT 598
601 TGGCTTTGAAATCAAAAGCTAGAGTGTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
599 TGGCTTTGTAATGCAAGGCTGAGAGTGTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 658
661 TGAATGCGTGAAGATATGGAAGACATCGATGGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
659 TGAATGCGTGAAGATATGGAAGACATCGATGGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 718
721 GACGCTCATGCAAGAAAGCTGGGAGAGCAAAAGAGATTAATACCTGATGATCAAGCC 780
719 GACGCTCATGCAAGAAAGCTGGGAGAGCAAAAGAGATTAATACCTGATGATCAAGCC 778
781 CT-AAACGATGCAACTAGTTGTTGGGCTTATTAATTAAGCTTGTGTAAGCAAGCTAAAGCGGTG 839
779 CTAAACGATGCAACTAGTTGTTGGGATTTCAATTTTCTCAGTAAAGCAAGCTAAAGCGGTG 838
840 AAGTTGACCGCTGGGAGAGTACGCTGCAAGATTAATACTCAAAAGAAATGACGGGAGCC 899
```

```
|||||
839 AAGTTGACCGCTGGGAGAGTACGCTGCAAGATTAATACTCAAAAGAAATGACGGGAGCC 898
900 CCGCAGAGCGGCTGATTAATGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 959
899 CCGCAGAGCGGCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 958
960 TGAATGATGCAAGAAATTTCTAGAGATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1016
959 TGAATGATGCAAGAAATTTCTAGAGATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1018
1017 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076
1019 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
1077 AACCTTGCATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1136
1079 AACCTTGCATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1133
1137 ACCGAGAGAGGCTGGGAGATGAGTCAAGTCTCATGAGCCCTTAATGAGGCTGAGGCTTCAAC 1196
1134 ACCGAGAGAGGCTGGGAGATGAGTCAAGTCTCATGAGCCCTTAATGAGGCTGAGGCTTCAAC 1193
1197 GTPAATCAATGAGCGGCTACAGAGGCTTGCACACCGCGAGGCGGAGCTAATCTCAGAAAG 1256
1194 GTCATCAATGAGCTGGTACAAAGGCTTGCACACCGCGAGGCGGAGCTAATCTCAGAAAG 1253
1257 CCGCTGCTGATGCTGGGAGTGGGAGTCTGCACTGCACTGCTGCACTGCTGCACTGCTGCACTG 1316
1254 CAGTCTGATGCTGGGAGTGGGAGTCTGCACTGCACTGCTGCACTGCTGCACTGCTGCACTG 1313
1317 TCGGAGATCAGATGCTGGGAGTGGGAGTCTGCACTGCACTGCTGCACTGCTGCACTGCTGCACT 1375
1314 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373
1376 ACCATGAGAGTGGGCTTCAACAGAGAGTGAATGATTAATGATTAATGATTAATGATTAATGAT 1435
1374 ACACCATGAGAGGCTTCTGCGAGAGTGAATGATTAATGATTAATGATTAATGATTAATGATTA 1433
1436 GG 1437
1434 GG 1435
```

Search completed: May 5, 2006, 14:06:52  
Job time : 1294 secs



SEQ ID NO 1  
LENGTH: 1508  
TYPE: DNA  
ORGANISM: Microsomona eutropha  
US-11-248-383-1

Query Match 84.2%; Score 1227; DB 18; Length 1508;  
Best Local Similarity 92.2%; Pred. No. 0;  
Matches 1349; Conservative 0; Mismatches 105; Indels 9; Gaps 5;

```
QY 1 ATTGAACGCTGGCGGCATCTTTACATGCAAGTCGAACGGCAGACCGATGCTTGCAT 60
DB 24 ATTGAACGCTGGCGGCATCTTTACATGCAAGTCGAACGGCAG--CGGGGCTTGGC 81
QY 61 CTGGTGGCAGTGGCGGACGGGTGAATGCAATCGAAAGTATCCAGAAAGGGGGGTA 120
DB 82 CTGGCGGCGAGTGGCGAAGGGGTGAATACATCGGAACGTGCTCTTGAAGGGGAATA 141
QY 121 AGGCATCGAAGATGCTAATACCGCATATCTCTAAGAGAAAGCGGGGATCGAAA 180
DB 142 AGCATCGAAGATGCTAATACCGCATAT--TTCTAAGAAAGAAAGCGGGATCGAAA 200
QY 181 GACCTTGGCCTTTGGAGCGCCGATGCTGATTAAGCTAGTTGTTGGGTTAAAGCCTAC 240
DB 201 GACCTTGGCCTTAAGAGAGCGCCGATGCTGATTAAGCTAGTTGTTGGGTTAAAGCCTAC 260
QY 241 CAAGCGCAGATCAGTATGTTGGTCTGAGAGACGACACGACCTGGGATCGAGACAG 300
DB 261 CAAGCGCAGATCAGTATGTTGGTCTGAGAGACGACACGACCTGGGATCGAGACAG 320
QY 301 GCCCGACCTCTACCGGAGGACGAGTGGGGAATTTGGACATGGGGCGAAGCCTGATC 360
DB 321 GCCCGACCTCTACCGGAGGACGAGTGGGGAATTTGGACATGGGGCGAAGCCTGATC 380
QY 361 CAGCAATGCGCGGTGAGTGAAGAGGCTTTCGGTGTGAAGCTCTTTCAGTCGAAGA 420
DB 381 CAGCAATGCGCGGTGAGTGAAGAGGCTTTCGGTGTGAAGCTCTTTCAGTCGAAGA 440
QY 421 AAAGGTTACGTTAATATCTGACCCATGACGGTATCGACAGAGAACACCGGCTAAC 480
DB 441 AAAGGTTACGTTAATATCTGACCCATGATTTATGACGGTATCGACAGAGAACACCGGCTAAC 500
QY 481 TACGTCGACGACCGCGGTAATAGCTAGGGTGAAGCGTTATCGGAATTTCTGGGCGT 540
DB 501 TACGTCGACGACCGCGGTAATAGCTAGGGTGAAGCGGTAATCGGAATTTCTGGGCGT 560
QY 541 AAAGGTCGCGACGCGCTTGTAACTGATGATGTAATCCCGGCGCTTAACCTGGGAAT 600
DB 561 AAAGGTCGCGACGCGCTTGTAACTGATGATGTAATCCCGGCGCTTAACCTGGGAAT 620
QY 601 TCGGTTGAACTACAAAGCTAGAGTGTGGCAGAGGGAGGTGAATTCATGTGTACAG 660
DB 621 TCGGTTGAACTACAAAGCTAGAGTGTGGCAGAGGGAGGTGAATTCATGTGTACAG 680
QY 661 TGAATATGCTAAGATATGAAGAACAATGATGGCGAAGCGACCTCCGCGGTTAACAT 720
DB 681 TGAATATGCTAAGATATGAAGAACAATGATGGCGAAGCGACCTCCGCGGTTAACAT 740
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGGTAGTCCAGCGC 780
DB 741 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGGTAGTCCAGCGC 800
QY 781 CTAAACGATGTCAACTAGTTGTGGGCTTATTA--GGCTGTGTAACGAAGCTAACGCGT 838
DB 801 CTAAACGATGTCAACTAGTTGTGGGCTTATTAAGATTGGTAACTAGCTAACGCGT 860
QY 839 GAAAGTTGACCGGCTGGGAGTACGTCGCAAGATTAAATCTAAAGAAATTGAACGGGAGC 898
DB 861 GAAAGTTGACCGGCTGGGAGTACGTCGCAAGATTAAATCTAAAGAAATTGAACGGGAGC 920
QY 899 CCGCACAAGCGGTGATATGTTGATTAATTCGATGCAACGCGAAGAAACCTTAACCTAAC 958
DB 921 CCGCACAAGCGGTGATATGTTGATTAATTCGATGCAACGCGAAGAAACCTTAACCTAAC 980
```

```
QY 959 TTGACATGTAGCAATTTTCTAGAGATGATTAAGTGC---TTCGGAAACGCTAACACAG 1015
DB 981 TTGACATCTTGGAAATCTAATGAGACATTAAGAGTCCCGCAAGAGACCAAGACACAG 1040
QY 1016 TGTGTGATGAGTGTGTGTCAGCTGTCGTGATGATGTTGGGTTAACTCCCGAACAGAGC 1075
DB 1041 TGTGTGATGAGTGTGTGTCAGCTGTCGTGATGATGTTGGGTTAACTCCCGAACAGAGC 1100
QY 1076 CAACCTTGTCTAATTAATGTCATCATTTT--GGTTGGGCACTTTAATGACATGCGGTGAC 1134
DB 1101 CAACCTTGTCTAATTAATGTCATCATTTAATAATGACACTTTAAGTGAATGCGGTGAC 1160
QY 1135 AAACCGGAGAAAGGTGGGATGACGTCAGAGTCTCATAGCCCTTAATGGGTAGGCTTAC 1194
DB 1161 AAACCGGAGAAAGGTGGGATGACGTCAGAGTCTCATAGCCCTTAATGGGTAGGCTTAC 1220
QY 1195 AGCTAATCAATGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAA 1254
DB 1221 AGCTAATCAATGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAA 1280
QY 1255 AGCGGTGTGATGTCGATTCGAGTCTGCAACTGCACTCCGTGAAGTGGAAATGCTACT 1314
DB 1281 AGCAGTGTGATGTCGAGTTCGAGTCTGCAACTGCACTCCGTGAAGTGGAAATGCTACT 1340
QY 1315 AATGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCA 1374
DB 1341 AATGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCA 1400
QY 1375 CACCATGGAGTGGGTTTACACAGAGAGGTAGTCAACCTGAAGAGAGCGCTTGCCA 1434
DB 1401 CACCATGGAGTGGGTTTACACAGAGAGGTAGTCAACCTGAAGAGAGCGCTTGCCA 1460
QY 1435 CGGTGAGATTCAATGACTGGGGTG 1457
DB 1461 CGGTGAGGTCAATGACTGGGGTG 1483
```

RESULT 2  
US-11-198-242-2  
Sequence 2, Application US/11198242  
Publication No. US20060035345A1  
GENERAL INFORMATION:  
APPLICANT: AJINOMOTO CO., INC.  
TITLE OF INVENTION: Process for the production of beta-amino acids by using acylase  
FILE REFERENCE: AB04037  
CURRENT APPLICATION NUMBER: US/11/198,242  
CURRENT FILING DATE: 2005-08-08  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 1522  
TYPE: DNA  
ORGANISM: Burkholderia sp.  
FEATURE:  
NAME/KEY: 16S rDNA  
LOCATION: (1)..(1522)  
OTHER INFORMATION:  
US-11-198-242-2

Query Match 78.9%; Score 1149.2; DB 18; Length 1522;  
Best Local Similarity 88.2%; Pred. No. 0;  
Matches 1287; Conservative 0; Mismatches 163; Indels 10; Gaps 3;

```
QY 1 ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGACCGATGCTTGCAT 60
DB 20 ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGACCGATGCTTGCAT 77
QY 61 CTGGTGGCAGTGGCGGACGGGTGAATGCAATCGAAAGTATCCAGAAAGGGGGGTA 120
DB 78 CTGGGCGGAGTGGGGAACGGGTGAGTAAATCATGGAACGCTGCTTGTAGTGGGGGATA 137
QY 121 AGGCATCGAAGATGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATCGAAA 180
```

```
Db      138 | G C C C G G A A A G C C G A T T A A T A C C G C A T A C G C T T A C G A G A A A G G G G G A A T T T A G 197
Qy      181 | G A C C T T G C G T T T T G A G C G G C C G A T G T C T A T T A G T A G T T G T G G G T A A A G G C T A C 240
Db      198 | G A C C T T C C G T A C A G G G G C C G C G A T G C A G A T T A G T A G T T G T G G G G T A A A G G C T A C 257
Qy      241 | C A A G G C A C G A T C A G T A G T T G T C T G A G A G A C A C A C C A C A C T G G G A C T A G A C A C G 300
Db      258 | C A A G G C A C G A T C T G T A G C T G T C T G A G A G A C A C A C C A C A C T G G G A C T A G A C A C G 317
Qy      301 | G C C C A A C T C T T A C G G A G G C A C A G T G G G A A T T T T G A C A A T G G G C G A A C C T G A T C 360
Db      318 | G C C C A A C T C T T A C G G A G G C A C A G T G G G A A T T T T G A C A A T G G G C G A A C C T G A T C 377
Qy      361 | C A G C A A T G C C G T G A G T A A A A G G C C T T C G G G T T G T A A A G C T T T C A G T C A G A A G A 420
Db      378 | C A G C A A T G C C G T G T G T A A A A G G C C T T C G G G T T G T A A A G C A C T T T T G T C G G A A A G A 437
Qy      421 | A A A G T T A C G G T A A T A T A T C T G A C C C A T A C G G T A T C G A C A G A A G A C A C G G C T A C 480
Db      438 | A A A C G C G T G T T A T A T A C C G T G G C G G A T A C G G T A C C G A A G A A T A A G C A C C G C T A C 497
Qy      481 | T A C G T C C A G A C C C G C G T A A T A C T A G G G T C A A C C T T A T C G A A T T A C T G G G C G T 540
Db      498 | T A C G T C C A G A C C C G C G T A A T A C T A G G G T C A A C C T T A T C G A A T T A C T G G G C G T 557
Qy      541 | A A A G G G G C G A G G C G G C T T T A A A G T C A G T G T A A T C C C G G G C T T A A C T G G A A T 600
Db      558 | A A A G C G T C C A G G C G G T C C G C T A A G A C A G T G A A A T C C C G G G C T T A A C T G G A A C 617
Qy      601 | T G C G T T T A A A C T A C A A A G C T A G A G T G G C A G A G G A G G T G A A T T C A T G T G T A C A G 660
Db      618 | T G C A T T T G T A C T G C G G G C T A G A G T A T G C A G A G G G G G A G A A T T C A C G T G T A C A G 677
Qy      661 | T G A A A T C G T A G A G A T T G A A A A C A T C C A T G C C A A G G C A G C C T C T G G T T A A C A C T 720
Db      678 | T G A A A T C G T A G A G A T T G A G A A T A C C G A T G G C A A G G A G C C C C T G G G C C A A A T C 737
Qy      721 | G A G C T A T G C A G A A A G C G T G G G A C A A C A G A T T A G A T C C T G T A G T C A C A G C C 780
Db      738 | G A G C T A T G C A G A A A G C G T G G G A C A A C A G A T T A G A T A C C T G T A G T C A C A G C C 797
Qy      781 | C T A A A C A T G C A C T A G T T G T T G G C C T T A T T A G C C T T G T A A C G A G C T A A C G C G T G A 840
Db      798 | C T A A A C A T G C A C T A G T T G T T G G G A T T C A T T C C T T A G T A C G T A A C G C G T G A 857
Qy      841 | A G T T A C C C C T G G G A G T A C G T C G A A G T T A A A C T C A A A G A A T T G A C G G G A C C 900
Db      858 | A G T T A C C C C T G G G A G T A C G T C G A A G T T A A A C T C A A A G A A T T G A C G G G A C C 917
Qy      901 | G C A C A C C G T G A T T A T G T G A T T A A T T G A T G C A A C G G A A A A C C T T A C C T A C C C T T 960
Db      918 | G C A C A C C G T G A T G A T G A T T A A T T C A T G C A A C G C A A A A A C C T T A C C T A C C C T T 977
Qy      961 | G A C A T G A C G A A T T T T C T A G A G A T A G A T T A G T G C - - - T T C G G A A C G C T A A C A G G T G 1017
Db      978 | G A C A T G A T G A A T C T G C T A G A G A G T G G A G T C C G A A A G G G A C C A T A A C A G A G T G 1037
Qy      1018 | C T G A T G G C T G T C G T C A G C T G T G T C T G A G A T G T T G G G T T A A G T C C C G A A C G A C G C A 1077
Db      1038 | C T G A T G G C T G T C G T C A G C T G T G T C T G A G A T G T T G G G T T A A G T C C C G A C A C A G C C A 1097
Qy      1078 | A C C C T T G C A T T A A T T G C A C A T T T G T T G G G A C C T T A A T G A C T G C C G T G A C A A 1137
Db      1098 | A C C C T T G T C C T A G T T G C T A C - - - - G A A G A G C A C T T A G G A G A C T G C C G T G A C A A 1152
Qy      1138 | C C G A G A A G G T G G G A T A G A G T C A A G T C T A T G C C C T T A T G G G T A G G G C T T C A C A C G 1197
Db      1153 | C C G A G A A G G T G G G A T A G A G T C A A G T C T A T G C C C T T A T G G G T A G G G C T T C A C A C G 1212
Qy      1198 | T A A T A C A A T G G C G C T A C A G A G G T T G C A A C C C G A G G G G A G C T A T C T A G A A A G C 1257
```

```
Db      1213 | T C A T A C A A T G T C G G A A C A G A G G T G C C C A C C C C G A G G G G A G C A A T C C A G A A A A C 1272
Qy      1258 | G G T C T A T C C G G A T C C G A A G T C G A A C T C G A C T C C G T A A T C G A A T C G T A A T 1317
Db      1273 | C G A T C G T A C C G A T C C G A T C T G C A C T C G A G T C G A A G C T G A A T C G T A G T A A T 1332
Qy      1318 | C G G A T C A G A T G T C C G G T A A T A C T T C C G G G T C T T G T A C A C A C C C C C G T C A C A C 1377
Db      1333 | C G G A T C A G A T C C G C G G T A A T A C T T C C G G G T C T T G T A C A C A C C C C C G T C A C A C 1392
Qy      1378 | C A T G G A G T G G G T T T C A C A A A G A C G A T G T C T A A C C G T A A G A G G G G C T T G C C A C G 1437
Db      1393 | C A T G G A G T G G G T T T A C C A A A G T G C T A G T C T A C C C A A G A G A G A G G T C A C A C G 1452
Qy      1438 | T G A G A T T C A T G A C T G G G G T G 1457
Db      1453 | T A G A T T A T G A C T G G G G T G 1472
```

```
RESULT 3
US-11-198-242-1
; Sequence 1, Application US/11198242
; Publication No. US20060035345A1
; GENERAL INFORMATION:
; APPLICANT: AJINOMOTO CO., INC.
; TITLE OF INVENTION: Process for the production of beta-amino acids by using acylase
; FILE REFERENCE: AB04037
; CURRENT APPLICATION NUMBER: US/11/198,242
; CURRENT FILING DATE: 2005-08-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1526
; TYPE: DNA
; ORGANISM: Variovorax sp.
; FEATURE:
; NAME/KEY: 16S rDNA
; LOCATION: (1)..(1526)
; OTHER INFORMATION:
US-11-198-242-1
```

Query Match 78.3%; Score 1141.2; DB 18; Length 1526;

Best Local Similarity 87.7%; Pred. No. 0;

Matches 1281; Conservative 0; Mismatches 173; Indels 6; Gaps 3;

```
Qy      1 | A T T G A A C G T G C G G A C T G C T T T A C A C A T C A A G T G A A C G C A C G A C G A T C T T G C A T 60
Db      20 | A T T G A A C G T G C G G A C T G C C T T A C A C A T C A A G T G A A C G C A C G C G G A G C A A T C - 77
Qy      61 | C T G G T G C G A T G C G G A C G G T G A G T A A T G C A T C G G A A G T A T C C A A G A G G G G G T A 120
Db      78 | C T G G C G C G A G T G C G A A C G G T G A G T A A T C G A A C G T G C C A A T C G T G G G G A T A 137
Qy      121 | A C G C A T G A A A G A T G T C T A A T A C C G C A T A C T C T A A G A G A A A G C A G G G A T C G A A 180
Db      138 | A C G A G A G A A G C T G T C T A A T A C C G A T A C G A T C A C G A T G A A A C A G A G G A T C C C A A 197
Qy      181 | G A C C T T G C G T T T T G A G C G G C C G A T G T C T A T T A G C T A G T T G T G G G T A A A G C C T A C 240
Db      198 | G A C C T T C C G A A T G A A G A C G G C C G A T G G C A G A T T A G T G G T G A G T A A A G C T C A C 257
Qy      241 | C A A G G C A C G A T C A G T A G T T G T T G A G A G A C A C A C C A C A C T G G G A C T A G A C A C G 300
Db      258 | C A A C C T T C A A T C T G A G C T G G T C T G A G A G A C A C A C C A C A C T G G A C T G A C A C G 317
Qy      301 | G C C A G A C T C C T A C G G A G G A G A C A G T G G G A A T T T T G A C A A T G G G C G A A C C T G A T C 360
Db      318 | G C C A G A C T C C T A C G G A G G A G A C A G T G G G A A T T T T G A C A A T G G G C G A A G C C T G A T C 377
Qy      361 | C A G C A A T G C C G C T G A G T A A A A G G C C T T C G G G T T A A A G C T T T C A G T C A G A A A G A 420
Db      378 | C A G C A A T G C C G C G T G A C A G A T A A A G C C T T C G G G T T A A A C T G C T T T T A C C G A A A G A 437
```



QY 421 AAAAGTTAGGTAATATCGTGAACCCATGACGGTATCGACAGAAAGACCGGCTAAC 480  
DB 438 AACGGCTCTTTCTTAATAAAGAGGGCTAATGACGGTACCGTAAGATAAGACCGGCTAAC 497  
QY 481 TAGGTCCAGACGCCCGGTAAATACGTAGGGTGCAGACGCTTAATCGGAATTAATCTGGGCGT 540  
DB 498 TAGGTCCAGACGCCCGGTAAATACGTAGGGTGCAGACGCTTAATCGGAATTAATCTGGGCGT 557  
QY 541 AAAAGGTGCGCAGCGCGCTTGTAGTCAAGATGTAAATCCCGGCGCTTAATCTGGGCGT 600  
DB 558 AAAAGGTGCGCAGCGCGGTGTAAAGACATTTGTAAATCCCGGCGCTTAATCTGGGCGT 617  
QY 601 TGGCTTTGAACTCAAAAGTGTAGGTGCGCAGAGGGAGGTGGAATTCATGTGTAGCAG 660  
DB 618 TGGCTGTGTGACGTGCTGAGTACGCGCAGAGGGAGGTGGAATTCGCGGTGTAGCAG 677  
QY 661 TGAATGCGGTAGATATGGAAGAAATCGATGGCGCAGACGCTCTGCGTTAACTACT 720  
DB 678 TGAATGCGGTAGATATGCGAGAGAACCGATGGCGCAGACGCTCTGCGGTGTACT 737  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAAGTACCTGTTAGTCCAGCGC 780  
DB 738 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAAGTACCTGTTAGTCCAGCGC 797  
QY 781 CTAAAGATGTCACTAGTTGTGGGCTTATTAAGCTTGTGTAGTACGAAGCTAACGCGTGA 840  
DB 798 CTAAAGATGTCACTAGTTGTGGGCTTATTAAGCTTGTGTAGTACGAAGCTAACGCGTGA 857  
QY 841 AGTTGACCGCTGGGAGTACGCTGCGCAAGATTAAACTCAAGAAATTGACGCGGAGCC 900  
DB 858 AGTTGACCGCTGGGAGTACGCGGCCCGCAAGTTGAATCAAGAAATTGACGCGGAGCC 917  
QY 901 GCAACAAGCGGTGATATGTGATTAATTCATGCAACGCGAAATTAATCTTACCTTCT 960  
DB 918 GCAACAAGCGGTGATATGTGATTAATTCATGCAACGCGAAATTAATCTTACCTTCT 977  
QY 961 GACATGTAGGAATTTTCTAGAGATAGATTAGTGTCT---TCGGGAAAGCGTAAACAGAGGT 1017  
DB 978 GACATGTAGGAATTTTCTAGAGATAGATTAGTGTCTGCGAAAGAAACCGTAAACAGAGGT 1037  
QY 1018 CTGCAATGCTGTCTGTCAGTCTGTGCTGATGATGTTGGTTAAAGTCCCGCAAGACGCA 1077  
DB 1038 CTGCAATGCTGTCTGTCAGTCTGTGCTGATGATGTTGGTTAAAGTCCCGCAAGACGCA 1097  
QY 1078 ACCCTTGTCTAATTAATGTCATCTTGTGGGCACTTTATGAGACTGCGGTGACAA 1137  
DB 1098 ACCCTTGTCTAATTAATGTCATCTTGTGGGCACTTTATGAGACTGCGGTGACAA 1156  
QY 1138 CCGAGAGAAAGTGGGATGACGTCAAGTCTCTCATGCGCTTAATGAGGTGGGCTTACACG 1197  
DB 1157 CCGAGAGAAAGTGGGATGACGTCAAGTCTCTCATGCGCTTAATGAGGTGGGCTTACACG 1216  
QY 1198 TAATTAATGAGCGCGTACAGAGAGTTGCCAAACCGCAGAGGGAGCTTAATCTCAGAAAG 1257  
DB 1217 TCAATTAATGAGCGCGTACAGAGAGTTGCCAAACCGCAGAGGGAGCTTAATCTCAGAAAG 1276  
QY 1258 GCGTGTAGTCCGAGATCGAGTCTGCACTGCTCGTAAGTCCGAATCGTGTAGTAT 1317  
DB 1277 CAGTGTAGTCCGAGATCGAGTCTGCACTGCTCGTAAGTCCGAATCGTGTAGTAT 1336  
QY 1318 CCGGATCAGCATGTGCGGTGAAATAGTTCGCGGCTTGTACACACCGCGCTGACAC 1377  
DB 1337 CCGGATCAGCAAGTGTCAAGTGTGATGTTCCGCGGCTTGTACACACCGCGCTGACAC 1396  
QY 1378 CATGGAGTGGGTTTACAGAGAGCAGGTAGTCTTAACCTTAAGAGAGGCGCTTGCCACG 1437  
DB 1397 CATGGAGTGGGTTTCTGCCAGAAAGTATGCTTAACCGCAAGAGGCGATTAACACG 1456  
QY 1438 TGAGATTCACTGAGGCTG 1457  
DB 1457 CAGGTTCTGTAAGTGGGCTG 1476

RESULT 4  
US-10-515-311-5  
; Sequence 5, Application US/10515311  
; Publication No. US20060030021A1  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Susan J  
; TITLE OF INVENTION: Novel Microorganisms and Uses therefor  
; FILE REFERENCE: 11752-006US1  
; CURRENT APPLICATION NUMBER: US/10/515,311  
; PRIOR FILING DATE: 2004-11-22  
; PRIOR APPLICATION NUMBER: PCT/NZ2003/000100  
; PRIOR FILING DATE: 2003-05-22  
; PRIOR APPLICATION NUMBER: NZ 519137  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 5  
; LENGTH: 1405  
; TYPE: DNA  
; ORGANISM: Acidivorax temperans  
US-10-515-311-5  
Query Match 74.9%; Score 1091.2; DB 11; Length 1405;  
Best Local Similarity 87.3%; Pred. No. 0;  
Matches 1234; Conservative 0; Mismatches 168; Indels 12; Gaps 3;  
QY 18 TCGTTTACATGCAAGTTCGAAACGGCAGACGATGCTTGCATCTGTTGGCGAGTGGCGG 77  
DB 1 TCGCTTACAGACGAGTGCAGCGGTAC---AGCTCTTGGATGCTGACGAGTGGCGA 56  
QY 78 ACCGGTGAATATGATCGGAACTATCCAGAAAGAGGGGGTAAAGCATTCGAAAGATGTG 137  
DB 57 ACCGGTGAATATATCATCGGAACGTGCCGATCGTGGGGGTAAGAAAGCGAAAGCTTTG 116  
QY 138 CTAAATACCGCATATATCTTAAGAGAGAAAGCAGGAGTGCAGAAAGCTTGGCTTTTGA 197  
DB 117 CTAAATACCGCATATATCTTAAGAGAGAAAGCAGGAGTGCAGAAAGCTTGGCTTTG 176  
QY 198 GGGCGCATGTCTGATTAGCTGTTGTGGGGTAAAGGCTTACCAAGGCAAGATCAGTA 257  
DB 177 GGGCGCATGTCTGATTAGCTGTTGTGGGGTAAAGGCTTACCAAGGCAAGATCAGTA 236  
QY 258 GTTGTCTGAGAGAGCAGACCACTGAGCTGAGACAGGCGCCAGACTCTTACCGG 317  
DB 237 GCTGTCTGAGAGAGCAGACCACTGAGCTGAGACAGGCGCCAGACTCTTACCGG 296  
QY 318 AGGCAAGCTGGGGAATTTTGAACAATGGCGCAAGCTTATCGAGCAATGCCGCTGAG 377  
DB 297 AGGCAAGCTGGGGAATTTTGAACAATGGCGCAAGCTTATCGAGCAATGCCGCTGAG 356  
QY 378 TGAAGAGGCTTGGGTTGTAAGCTCTTTCAGTCGAGAAAGAAAGTTAGGTAAATA 437  
DB 357 GATGAAAGGCTTGGGTTGTAAGCTCTTTCAGTCGAGAAAGAAAGTTAGGTAAATA 416  
QY 438 ATCGTGAACGATGACGTTATCGACAGAAAGCAGCGGTAACTAGTCCAGACGCGC 497  
DB 417 CCGGAGTCTATGACGATTCCTTAAGAAATAGACACCGGCTAATCTGTCAGACGCGC 476  
QY 498 GGTAAATACGTAGGTTGCAAGCGTTAATCGAAATTAATGAGGCTGAGGCGGCGG 557  
DB 477 GGTAAATACGTAGGTTGCAAGCGTTAATCGAAATTAATGAGGCTGAGGCGGCGG 536  
QY 558 CCTTGTATGATCAGATGTGAAATCCCGGCGCTTAACCTGGGAAATTCGTTTGAACATCA 617  
DB 537 TTAATTAACAGATGTGAAATCCCGGCGCTTAACCTGGGAAATTCGTTTGAACATCA 596  
QY 618 AGCTAGAGTGGCAGAGGAGGTGGAATTCATGTGTAGCAGTGAATGCTAGAGATA 677  
DB 597 AGCTAGAGTGGCAGAGGAGGTGGAATTCGCTGTGTAGCAGTGAATGCTAGAGATA 656  
QY 678 TGAAGAAATCATGATGCGAAGAGCAGCTTCTGAGTTAACTAGACGCTCATGACGAA 737  
DB 657 CCGAGAGAACCGATGCGAAGAGCAATCCCTGGGCTGTACTGACGCTCATGACGAA 716



QY 738 GCGTGGGGAGGAGAAACAGGATTATAGATACCTGGTAGTCCAGCCGCTTAAAGATGTCAACTTA 797  
 Db 717 GCGTGGGGAGGAGAAACAGGATTATAGATACCTGGTAGTCCAGCCGCTTAAAGATGTCAACTG 776  
 QY 798 GTTGTGTGGGCTTATTTAAGGCTTGTATACGAGCTTAAACGCGTGAAGTTGAACCGCTTGGGGA 857  
 Db 777 GTTGTGTGGGCTTATTTAAGGCTTGTATACGAGCTTAAACGCGTGAAGTTGAACCGCTTGGGGA 836  
 QY 858 GTACCGGTGCGAAGATTAAACTCMAAGAAATTGACGGGGGACCCGGCAACAGCGGTGGAATTA 917  
 Db 837 GTACCGCGCGCAAGGTTGAACCTCAAGAAATTGACGGGGGACCCGGCAACAGCGGTGGAATGA 896  
 QY 918 TGTGGAATTAAATTCGATGACGACGCGAAAACTTTACCTACCTTGAATGATGACGAATTTT 977  
 Db 897 TGTGGAATTAAATTCGATGACGACGCGAAAACTTTACCTTGAATGATGATGACGAATTCCT 956  
 QY 978 CTAGAGATAGATTAGTGCT--TCGGGACGCTTACACAGGTCTGTCATGCTGTGCTCA 1034  
 Db 957 TTAGAGATAGAGAGAGTGCTCGAAAGAGACCCGTAAACAGGTGCTGCATGCTGTGCTCA 1016  
 QY 1035 GCTCGTGTCTGATGATGTGGGTTAAGTCCCGCAAGAGCGCAACCTTGTCAATTATG 1094  
 Db 1017 GCTCGTGTCTGATGATGTGGGTTAAGTCCCGCAAGAGCGCAACCTTGTCAATTATG 1076  
 QY 1095 CCATCATTTGTTGGGCACTTTAATAGACTGCGCGGTGACAAACCGGAGGAAAGTGGGGA 1154  
 Db 1077 C-----TAGAGAAAGGCACTTATAGGAGCTGCGGTGACAAACCGGAGGAAAGTGGGGA 1131  
 QY 1155 TGACGCTCAAGTCTTCAATGAGCCCTTATGGGTAGGGCTTACACGTAATACATGCGCGCTA 1214  
 Db 1132 TGACGCTCAAGTCTTCAATGAGCCCTTATGGGTAGGGGTCTACACGCTCATATGCTGTGTA 1191  
 QY 1215 CAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAGAAAGGCGGTGTAATCCGGATC 1274  
 Db 1192 CAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCCATTAAGCAAGTCGATGTCGGATC 1251  
 QY 1275 GGAAGTCTGCAACTCCGACTCCGTGAAGTCGGAAATCGCTAGTAATCGCGGATCAGCATGTCG 1334  
 Db 1252 GCAAGTCTGCAACTCCGACTCCGTGAAGTCGGAAATCGCTAGTAATCGCGGATCAGCATGTCG 1311  
 QY 1335 CGGTGAATACGTTCCCGGGCTTTGTACACACCGCCGCTCACACCATGTGGAGTGGCTTTC 1394  
 Db 1312 CGGTGAATACGTTCCCGGGCTTTGTACACACCGCCGCTCACACCATGTGGAGCGGGTTCG 1371  
 QY 1395 CCAAGAACGAGTATGACTTAAACCGTAAGAGGGGCGC 1428  
 Db 1372 CCAAGAACGAGTATGACTTAAACCGTAAGAGGGGCGC 1405  
 RESULT 5  
 US-10-831-286A-48683  
 / Sequence 48683, Application US/10831286A  
 / Publication No. US20060046246A1  
 GENERAL INFORMATION:  
 / APPLICANT: ZENG, QIANDONG  
 / APPLICANT: CHATELIER, SONIA  
 / APPLICANT: MOIR, DONALD T.  
 / APPLICANT: LACROIX, BRUNA  
 / APPLICANT: CHILDRESS, DARRELL  
 / TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
 / FILE REFERENCE: 032796-174.001  
 / CURRENT APPLICATION NUMBER: US/10/831,286A  
 / PRIOR FILING DATE: 2004-04-26  
 / PRIOR APPLICATION NUMBER: 60/464,955  
 / NUMBER OF SEQ ID NOS: 48788  
 / SOFTWARE: PatentIn version 3.2  
 / SEQ ID NO 48683  
 / LENGTH: 1493  
 / TYPE: DNA  
 / ORGANISM: Pantoea dispersa  
 US-10-831-286A-48683

Query Match	70.1%	Score 10.1	2:DB 11:	Length 1493:
Local Similarity	82.9%	Pred. No. 5	3e-307	
Matches 1214:	Conservative	0:	Mismatches 243:	Indels 7:
Gaps				4:
QY	1	ATTGAACGCTGCGCGCATCTTTACACATGCAGATCGAACGCGACGACGAGTCTTG---	57	
Db	1	ATTGAACGCTGCGCGCATCTTTACACATGCAGATCGAACGCGACGACGAGTCTTG	60	
QY	58	-CATCTGTGTGGCGATGTGGCGAAGCGGTGAGTAAATGCAATCGGAACGTAATCCAGAAAGAGGGG	116	
Db	61	TCTTTGGGTGGCGAAGTGGCGAAGCGGTGAGTAAATGTCTGGAAACTGCCCCGATGTAGGGG	120	
QY	117	GGTAACGATCCGAAGAATGTCTAAATCCGACATATCTTAAGAGAGAAAGACGAGGATC	176	
Db	121	GATTAACTAAGTAAGAAACGTAAGTACTTAATACCGCATACGTCCGACAGACCAAAAGTGGGGACC	180	
QY	177	GAAAGACCTTGCGCTTTTGGAGCGCGCGATGTCTGATTTAGCTAGTTGTGGGGTAAAGGC	236	
Db	181	TTGGGGCTTCACACATCCGATGTGGCCAGATGGATTTAGCTAGTGTGGGGTAAATGGC	240	
QY	237	CTAACCAAGCGACGATCAGTAGTTGGTCTGAGAGAGACGACGACGACCACTGGGACCTGAGA	296	
Db	241	TCACTTAGGCGACGATCTCCAGCTGAGTCTGAGAGAGATGACACACCACTGGAACCTGAGA	300	
QY	297	CAGGCGCCACATCTCTACCGGGAAGGACAGAGTGGGGAAATTTGGACAATGGCGCAAGCCT	356	
Db	301	CAGGCTCAACATCTCTACCGGGAAGGACAGAGTGGGGAAATTTGGACAATGGCGCAAGCCT	360	
QY	357	GATCCAGACATGCGCGCTGAGTGAAGAGAGGCTTCGGGTTGTAAAGCTCTTTCAATGTCAG	416	
Db	361	GATGCAACCATGCGCGCTGATGAAGAGAGGCTTCGGGTTGTAAAGTACTTTACGCGGGG	420	
QY	417	AAGAAAAGTTACGGTAAATTAATCGTGAACCCATGACCGTATTCGACAGAGAAAGCACCGGC	476	
Db	421	AGGAAGCGGTGAGGTTAATTAATCCTTCCGATGACGTTACCGCAGAAAGACACCGGC	480	
QY	477	TAACTACGTGCGACGACCGCCGGTAAATGCTAGGGGTCCAAAGCCTTAATCCGAATTAATCG	536	
Db	481	TAACTCCGTGCGACGACCGCCGGTAAATGCGAGGGGTCCAAAGCCTTAATCCGAATTAATCG	540	
QY	537	GCGTAAAGGGATGCGCAGCGCGCCTTGTAAGTCAGATGTGAATATCCCGGGCTTAATCCCTG	596	
Db	541	GCGTAAAGCGACCGACGGCGGTCTGTAAGTCAGATGTGAATATCCCGGGCTTAATCCCTG	600	
QY	597	GAATTGCGTTGAAACTTACAAAGCTAGAGTGTGCGACAGGGAAGGTGAATTCATGTGTA	656	
Db	601	GAACTGCAATTTGAAACTGCGACGGCTTGAATCTGTAGAGGGGGGTGAATTCACAGGTGTA	660	
QY	657	GCAGTGAATATCGTAGAGATATGGAAGAACATGATGTCGCAAGGCAAGCCTCTCGGGTTAA	716	
Db	661	GCAGTGAATATCGTAGAGATCTGGAAGAAATACCGGTGCGCAAGGCGGCCCTCTGGAAGAA	720	
QY	717	CACGTAGCGTCAATCCAGAAAGCGTGGGGAGCAAAACGAAATTAGATACCTCGTAGTGTCCA	776	
Db	721	GACTAGCGCTCAGGTGCGAAAGCCTGGGGAGCAAAACGAAATTAGATACCTCGTAGTGTCCA	780	
QY	777	CGACCTAAACGATGTCAACTAGTTGTGGGCTTATTAAGGCTTGG-TAAAGAACCTAACG	835	
Db	781	CGCGCTAAACGATGTCAACTTGAAGGTTGTGCCCTTAGAGGTGTGGCTTCGGAAGCTAACG	840	
QY	836	CGTGAAGTGAACGCGCTGGGGAGTACGSGTGCAGAAAGATTTAAACTCAAAAGAAATTTGACGG	895	
Db	841	CGTTAAGTGCACCGCTGGGGAGTACGCGCGCAAGGTTTAAAACTCAATATGATTTGACCGG	900	
QY	896	GACCCGCAAGACGCGGTGATTAATGTGATTTAAATTCGATGCAAGCGCAAAAACCTTAACCTA	955	
Db	901	GCGCCGCAAGACGCGTGAAGCATGTGTTTAATTCATGCAACCGCAAGAACTTAACTCTG	960	
QY	956	CCCTTGACATGACGAATTTCTAGAGATAGATTAGTGTG-CTTGCGGAACGCTAACACAG	1014	
Db	961	GCCTTGACATCAAGAACTTAGACAGAGATCTTTGGTGTGCTTGGGGAACCTTAGACAG	1020	

OY	1015	GTGCTGCATGGCTGTCGTCGACCTCGTGTGTCGAGATGTTGGGTTAAATGTCGCGAAGCAGC	1074
Db	1021	GTGCTGCATGGCTGTCGTCGACCTCGTGTGGAATGTTGGGTTAAATGTCGCGAAGCAGC	1080
OY	1075	GCAACCCCTGTGCATTAATTGGCCATC-ATTGGTTGGGCACTTTATATAGACTGCGGATGA	1133
Db	1081	GCAACCCCTTATTCCTTTGTGTGTCAGCGGTTCCGGCCGGGAATCAAAAGAGACTGCGCGGTGA	1140
OY	1134	CAAAACCGAGAGAGGTGGGGGATGACGTCAATGCTCCATATGCCCCCTTATGGGTGAAGGCTTCA	1193
Db	1141	TAAACCGGAGAGAGGTGGGAGTGAAGTAAATCATCAATGAGCCCTTTACGGCCAGGACTCA	1200
OY	1194	CACGTAATACATGGCGCGCTACAGAGGGTTGCCAAACCGCGAGGGGAGCTAATCTCAGA	1253
Db	1201	CACGTGCTCAATGGCCGCTACAAAGAAAGAACGACTGCGAGAGCAAGCCGACTCATTA	1260
OY	1254	AAGCGCGTCGTAGTCCGGATCGAGTCTGCAACTGCACCTCCGTGAAGTCGAAATCGCTAG	1313
Db	1261	AAGTCCGTCGTAGTCCGGATTTGGAGTCTGCAACTGCATCCATGAAAGTCGAAATCGCTAG	1320
OY	1314	TAAATCGCGATCAGCATGTCCGCGGATTAAGTTCCGGGGCTTTGATACACCGCCCGTCTC	1373
Db	1321	TAAATGTAGATCAGATGCTACGAGGAATAGCTTCCGGGGCTTTGATACACCGCCCGTCTC	1380
OY	1374	ACACCATGGAGTGGGTTTCAACGAAGACAGGTACTTAAACCGTAAGAGAGGGCGCTTGCC	1433
Db	1381	ACACCATGGAGTGGGTTTCAAAAGAAAGTAAGGTACTTAACTTTCGGAGAGGGCGCTTAACC	1440
OY	1434	ACGGTGAATTCATGATCTGGGGGTG	1457
Db	1441	ACTTTGTATTCATGACTGGGGGTG	1464

```

RESULT 6
US-10-513-639-1
; Sequence 1, Application US/10513639
; Publication No. US2006001051A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yicheng
; APPLICANT: Chen, Yancheng
; APPLICANT: Li, Fengmei
; APPLICANT: Tian, Zhexian
; APPLICANT: Lin, Min
; APPLICANT: Wang, Yiping
; TITLE OF INVENTION: NOVEL GLYPHOSATE-TOLERANT
; TITLE OF INVENTION: 5-ENOILPRUVYLSHIMIMATE-3-PHOSPHATE SYNTHASE AND THE GENE
; TITLE OF INVENTION: ENCODING THE SAME
; FILE REFERENCE: 18495-002US1
; CURRENT APPLICATION NUMBER: US/10/513,639
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: CN 02117991.3
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: CN 02117647.7
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: PCT/CN02/00539
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas putida P.P4G-1
US-10-513-639-1

```

Query Match	70.0%;	Score 1020.2;	DB 11;	Length 1501;
Best Local Similarity	83.0%;	Pred. No. 1.1e-306;		
Matches 1212;	Conservative	0;	Mismatches 233;	Indels 6; Gaps 4

Oy 1 ATTGAACGCTGGGCGGCATGCTTTACACATGCAAGTGCAGACGCGACGACCGATGCTTGCAT 60  
 Db 21 ATTGAACGCTGGGCGGCAGGCTTAAACATGCAAGTGCAGCGCGATGAGAAAGCTTGCCT 80  
 Oy 61 CTGGTGGCGAGTGGCGGACGGGTGAAGTAAATGCATCGGAACGTATCCAGAAAGAGGGGGTA 120

[illegible]

Db 1099 C C C T T G T C T T A G T T A C C A G C A C G T A T G T G G G C A C T T A G A G A C T G C C G T G C A A 1155  
Gy 1137 A C C G A G A A G T G G G G A T G A C G T C A G T C C T C A T G G C C C T T A T G G G A T A G G C C T T C A C A C 1191

Db 1159 ACCGAGAAAGTGGGATGGCGTCAAGTATCATAGGCCCTTAAGCGCTGGGCTACAC 1218  
Qy 1197 GTAATCAATGCGCGGTACAGAGGTTGCCAACCCGAGAGGAGCTAATCTCAGAAAG 1256  
Db 1219 GTGCTCAATGTCGGTACAGAGGTTGCCAACCCGAGAGGAGCTAATCTCAGAAAA 1278  
Qy 1257 CGGCTGTAGTCCGGATCGGAGTCTGCACTCCGACTCCGGAAGTCCGTAATCCCTGTA 1316  
Db 1279 CCATCGTAGTCCGGATCGGAGTCTGCACTCCGACTCCGGAAGTCCGTAATCCCTGTA 1338  
Qy 1317 TCCGATCGAGTATGTCGGGTGATGATGTTCCGGGCTTGTACACAGCCCGCTGACA 1376  
Db 1339 TCCGATCGAGTATGTCGGGTGATGATGTTCCGGGCTTGTACACAGCCCGCTGACA 1398  
Qy 1377 CCATGGAGTGGGTTTCCAGCAAGACAGTATCTAACCCTAAGAGAGGCGCTTGCACG 1436  
Db 1399 CCAATGGAGTGGGTTTCCAGCAAGAGTATGCTAATCTTCCGAGAGAGGCTTACACG 1458  
Qy 1437 GTGAGATTCAATGACTGGGGTG 1457  
Db 1459 GTGAGATTCAATGACTGGGGTG 1479

RESULT 7  
US-10-831-286A-48685  
; Sequence 48685, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHIDRESS, DARRELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-1174.001  
; CURRENT APPLICATION NUMBER: US/10/831, 286A  
; PRIOR FILING DATE: 2004-04-26  
; PRIOR FILING DATE: 2003-04-24  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 48685  
; LENGTH: 1508  
; TYPE: DNA  
; ORGANISM: Providencia rustigianii  
US-10-831-286A-48685

Query Match 70.0%; Score 1020.2; DB 11; Length 1508;

Best Local Similarity 83.3%; Pred. No. 1.1e-306;

Matches 1219; Conservative 0; Mismatches 238; Indels 6; Gaps 5;

Qy 1 ATTGAACGCTGGCGGATGCTTTACACATGCAATCGCAACGGCAGCAGG-GATGCTTGCA 59  
Db 11 ATTGAACGCTGGCGGCGGCTTAACACATGCAAGTCGACGCTTAACAGAAAGCTTGCT 70  
Qy 60 TCTGG-TGGGAGTGGCGGAGCGGGTGAATATGATCGGAAAGTATCCAGAAAGAGGSGG 118  
Db 71 TCTGCTGACAGCGCGGAGCGGGTGAATATGATGAGGATCTCCCGATAGAGGGGA 130  
Qy 119 TAAAGCATGAAGAATGTGCTAATACCGCATATACTCTAAGAGAGAAAGCAGGGGATGA 178  
Db 131 TAACTACTGGAACGCTAGCTAATACCGCATATCTCTCAGAGCAAAAGCAGGGAACTT 190  
Qy 179 AAGACCTTGGCGCTTTTGAAGCGGCGGATGTCTGATTTAGCTAATGTTGGGTTAAGGCT 238  
Db 191 CGGTCTTGGCGCTATCGATGAACCCATATGGGATTACTAATGTGGGTTAATGGCTC 250  
Qy 239 ACCAAGGAGAGCATGATGATGATCTGAGAGAGCAGCAGGCACTGGGATGAGACA 298  
Db 251 ACCAAGGAGAGCATGATGATGATCTGAGAGAGATGATGAGCAGCAGCTGGGATGAGACA 310  
Qy 299 CGGCGCAGATCTCTACGAGAGGACAGCATGTGGGAAATTTTGAACAATGGCGGCAAGCTTGA 358

Db 311 CGGCCAGACTCTCTACGGAGAGCAGCATGTGGGAAATATTGCACAATGGCGCAGACCTTGA 370  
Qy 359 TCCAGCAATGCGCGGTGAGTGAAGAGGCTTCCGGTGTGAAGCTCTTTCAGTCCAGAA 418  
Db 371 TGCAGCATTGCGCGGTGATGAAGAGGCTTGAAGGTTGAAGTACTTTCAGTTGGAG 430  
Qy 419 GAAAAGTTACGGTAAATATATCGTGAACCATGACGGTATCGACAGAAAGACCGGCTTA 478  
Db 431 GAAGGCGTTGATGCTAATATCGCAACGATGACGTTACCAACAGAAAGACCGGCTTA 490  
Qy 479 ACTACGTGCGAGAGCGCGGTATATCGTAGGCTGCAAGCGTTAATCGGAATTAATCGGC 538  
Db 491 ACTCGGTGCGAGAGCGCGGTATATCGAGGCTGCAAGCGTTAATCGGAATTAATCGGC 550  
Qy 539 GTAAAGGTGCGAGCGCGCTTGTAAATGATGTAATGTAATCCCGGGGCTTAACCTGGGA 598  
Db 551 GTAAACGCGACGAGCGCGGTGATTAATGATGTAATGTAATCCCGGGGCTTAACCTGGGA 610  
Qy 599 ATTGCGTTGAACCTCAAAAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGC 658  
Db 611 ATGCAATCTAAGACTGTGCTAGCTAGTCTTGTAGAGGGGGTGAATTCATGTGTAGC 670  
Qy 659 AGTGAATGCGTAGAGTATGAAGAAATCATGATGGCGAAGGAGCCTCTGGGTTAACA 718  
Db 671 GTGGAATGCGTAGAGTATGAAGAAATCATGATGGCGAAGGCGCCTCTGGACAAAGA 730  
Qy 719 CTGACGCTCATGCAACGAAAGCGTGGGAGAGCAACAGATTAGTACCCTGTTAGTCCAGC 778  
Db 731 CTGACGCTCATGCGCAAGCGTGGGAGAGCAACAGATTAGTACCCTGTTAGTCCAGC 790  
Qy 779 CCTTAAACGATGTCATAGTGTGGGCTTATTAAGGCTTGG-TAAGAAAGCTAAGCGC 837  
Db 791 CTGTAACGATGTCATAGTGTGGGCTTATTAAGGCTTGG-TAAGAAAGCTAAGCGC 850  
Qy 838 TGAAGTTGACCGGCTGGGAGTACGCTGCGAAGATTAACTCAAGAAATTAAGCGGGA 897  
Db 851 TTAATATGACCGGCTGGGAGTACGCGCGCAAGGTTAACTCAAGAAATTAAGCGGGA 910  
Qy 898 CCGGACAGCGGCTGATTAATGAGATTAATGATGTCACACGGAAACCTTAACCTACC 957  
Db 911 CCGGACAGCGGCTGATGATGATTAATGATGTCACACGGAAACCTTAACCTACC 970  
Qy 958 CTGACATGTAAGGAATTTCTAGAGTATGATTAAGTCTTCCGGAACGCTAACAAGT 1016  
Db 971 CTGACATGTAAGGAATTTCTAGAGTATGATTAAGTCTTCCGGAACGCTAACAAGT 1030  
Qy 1017 GCTGATGCTGCTGCTGATGCTGCTGCTGATGATGCTGCTGATGATGCTGCTGATGATG 1076  
Db 1031 GCTGATGCTGCTGCTGATGCTGCTGCTGATGATGCTGCTGATGATGCTGCTGATGATG 1090  
Qy 1077 AACCTTGTCAATTAATGCAATCA-TTGGTGGGCACTTTAATGAGACTGCCGCTGAC 1134  
Db 1091 AACCTTATCTTTGTTGTCAGACGTCATGCTGGAATCTCAAGAGAGACTGCCGCTGAT 1150  
Qy 1135 AAACCGAGAGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1194  
Db 1151 AAACCGAGAGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1210  
Qy 1195 ACCTAATACATGCGGCTGACAGAGGCTGCAACCGCGAGAGGAGCTAATCTCAGAA 1254  
Db 1211 ACCTGCTACATGCGGCTGATACAGAGGCTGCAACCGCGAGAGGAGCTAATCTCAGAA 1270  
Qy 1255 AGCGGCTGATGCTGCGATCGGATCTGCACTGCACTCCGCTGAGATGCGAATGCTAGT 1314  
Db 1271 AGTACGTGATGCTGCGATGCGAATCTGCACTGCACTCCGCTGAGATGCGAATGCTAGT 1330  
Qy 1315 AATGCGGATGACATGCTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1374  
Db 1331 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1390  
Qy 1375 CACCATGGGAATGGGTTTCAACAGAGAGGATGATTAACCTGAAGAGGCGCTTGGCA 1434  
Db 1391 CACCATGGGAATGGGTTTCAACAGAGAGGATGATTAACCTTGGGAGGCGCTTGGCA 1450



FILE REFERENCE: 032796-174.001  
CURRENT APPLICATION NUMBER: US/10/831.286A  
CURRENT FILING DATE: 2004-04-26  
PRIOR APPLICATION NUMBER: 60/464.955  
PRIOR FILING DATE: 2003-04-24  
NUMBER OF SEQ ID NOS: 48788  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 48681  
LENGTH: 1494  
TYPE: DNA  
ORGANISM: *Pantoea dispersa*  
US-10-831-286A-48681

Query Match 70.0%; Score 1019.6; DB 11; Length 1494;  
Best Local Similarity 82.9%; Pred. No. 1,76-306;  
Matches 1213; Conservative 0; Mismatches 244; Indels 7; Gaps 4;

QY 1 ATTGAACGGCTGGGGGCGATGCTTTACACATGCAAGTCGAAACGCGACGACGATGCTTGG--- 57  
DB 1 ATTGAACGGCTGGGGGCGAGGCTTAACATGCAAGTCGAAACGCGACGACGACGACGATGCTTGC 60  
QY 58 -CATCTGGTGGCGAGTGGCGGACGCGGTGAATGCAATCGAAACGTAATCGAAGAGGAG 116  
DB 61 TCTTTGGGTGGCGAGTGGCGGACGCGGTGAATGCTGGGAAACGTCGCGATGAGGAGG 120  
QY 117 GGTAACGATCGAAAGATGTGCTAATCCGCAATTAATCTTAAGAGAGAAAGCAGGAGATC 176  
DB 121 GATTAACATCGAAACGCTAGCTAATCCGCAATTAATCCGCAATTAATCGAAAGAGGAGACC 180  
QY 177 GAAAGACCTTGGCGCTTTTGAAGCGCGGATGCTGATTAAGTGTGGGGGTAAAGGC 236  
DB 181 TTGGGGCTCAACACATCGAGATGTGCCAGATGGGATTAAGTAAAGTGGGTAATGGC 240  
QY 237 CTACCAAGCGGACGATGATGTTGCTGAGAGAGACGACGACGACGACGACGACGACGACGACG 296  
DB 241 TCACCTGGGAGGACGATCCCTAGCTGTCTGAGAGAGAGACGACGACGACGACGACGACGACG 300  
QY 297 CACGGCCAGACTCTTAACGAGAGGACGACGAGTGGGAAATTTTGAACAATGGCGCAGACCT 356  
DB 301 CACGGCTCAACATCTTAACGAGAGGACGACGAGTGGGAAATTTTGAACAATGGCGCAGACCT 360  
QY 357 GATTCAGCAATGCCGCTGAGTGAAGAAAGGCTTCGGGTTGTAAGCTCTTTCAGTGGAG 416  
DB 361 GATGCAACCAATGCCGCTGAGTGAAGAAAGGCTTCGGGTTGTAAGCTCTTTCAGTGGAG 420  
QY 417 AAGAAAGGTTACGGTAATTAATGCTGACCCATGACGCTAATCGAAGAGAGACCGGC 476  
DB 421 AAGAAAGGTTGAGTTAATTAATCTCCGCAATTAATCGTAACCCGCAAGAGAGACCGGC 480  
QY 477 TAACTACGTCGACAGCGCGGTAATTAATGAGGAGTGAACGTTAATCGAATTAATCGG 536  
DB 481 TAACTCCGTCGACAGCGCGGTAATTAATGAGGAGTGAACGTTAATCGAATTAATCGG 540  
QY 537 GCGTAAAGGTGGCGAGCGCGCTTTGTAAGTCAAGTGAATATCCCGGGCTTAACTCGG 596  
DB 541 GCGTAAAGCGCACCGCGCGCTTTGTAAGTCAAGTGAATATCCCGGGCTTAACTCGG 600  
QY 597 GAATTTGGTTGAATACTAACAAGCTAGAGTGGCAGAGGAGGAGGTAATCCATGTGTA 656  
DB 601 GAATCTGATTTGAATACTGAGGCTTGAAGTCTCGTAAGGGGGGTAATCCAGGTGTA 660  
QY 657 GCACTGAATGCTAGAGATATGGAAGAACATCGATGCGAAGGACGCTCTGGGTTAA 716  
DB 661 GCGGTGAATGCTAGAGATATGGAAGAACATCGATGCGAAGGACGCTCTGGGTTAA 720  
QY 717 CACTGACGCTATGACGAAAGCGTGGGAGCAAAACAGATTAATGATACCTGCTGATGCTCA 776  
DB 721 GACTGACGCTATGAGGAGCAAAAGCGTGGGAGCAAAACAGATTAATGATACCTGCTGATGCTCA 780  
QY 777 GCGCTTAATGATGCACTAGATGTTGGGCTTAATTAGGCTTGG--TAAAGAGCTAAGC 835  
DB 781 GCGCTTAATGATGCACTAGATGTTGGGCTTGGGCTTGGGCTTGGGAGCTAAGC 840

QY 836 CGTGAAGTTGACCGCCTGGGAGATACGCTCGCAAGATTAAACTCAAGGAATTGACGG 895  
DB 841 CGTTAAGTGCACCGCTGGGAGATACGCGCCGCAAGGTTAAACTCAAAAGAAATGACGG 900  
QY 896 GACCCGCAACAGCGGTGATTAATGAGATTAATGATGCAACGCGAAACCTTACCTTA 955  
DB 901 GCGCCGCAACAGCGGTGATTAATGAGATTAATGATGCAACGCGAAACCTTACCTTA 960  
QY 956 CCCTGACATGTAGCAATTTTCTAGATTAATGATG--CTTGGGAAACGCTTAACAG 1014  
DB 961 GCCTTACATTCAGAGAACTTGAAGAGATGCTTGTGCTTGGGAACTCTGAGACAG 1020  
QY 1015 GTGCTGATGCTGTGTCGACGCTGTGTCTGAGATGTTGGTTAAGTCCCGACAGAGC 1074  
DB 1021 GTGCTGATGCTGTGTCGACGCTGTGTCTGAGATGTTGGTTAAGTCCCGACAGAGC 1080  
QY 1075 GCAACCTTGTCTTAATTTGCCATC-ATTGTTGGGCACTTAATGAGCTGCGGTGA 1133  
DB 1081 GCAACCTTATCCTTTGTGCGACGCGTTTCGCGGGAACCTAAAGAGACTGCGGTGA 1140  
QY 1134 CAACCGGAGAGAGTGGGAGTGAAGTCAAGTCTCATGGCCTTAATGGGTAGGGCTTCA 1193  
DB 1141 TAAACCGAGAGAGTGGGAGTGAAGTCAAGTCTCATGGCCTTAATGGGTAGGGCTTCA 1200  
QY 1194 CAGTAATACATGAGCGCGCTACAGAGGTTGCCAACCCCGAGAGGAGGAGCTTAATTCAGA 1253  
DB 1201 CAGTGTCTACATGAGCGCATACAAAGAGAGAGACGCTCGGAGAGAGGAGGAGCTTATTA 1260  
QY 1254 AAGCGGTGCTAGTCCGGATCGGAGTCTGCACTGCACTCGTGAAGTGGGAATCGGTAG 1313  
DB 1261 AAGTGGTGTAGTCCGGATCGGAGTCTGCACTGCACTCGTGAAGTGGGAATCGGTAG 1320  
QY 1314 TAAATCGGATGACGATGTGCGCGTGAATTCGTTCCGGGCTTTGTAACACGCGCGCTC 1373  
DB 1321 TAAATCGGATGACGATGTGCGCGTGAATTCGTTCCGGGCTTTGTAACACGCGCGCTC 1380  
QY 1374 ACACCAATGGAGTGGGTTTCAACCAAGCAGTAGTCTTAACCGTAAGAGAGGCGCTTACC 1433  
DB 1381 ACACCAATGGAGTGGGTTTCAACCAAGCAGTAGTCTTAACCGTAAGAGAGGCGCTTACC 1440  
QY 1434 ACGGTGATTCATGACTGGGGTG 1457  
DB 1441 ACTTTGATTCATGACTGGGGTG 1464

RESULT 10  
US-10-831-286A-48677  
Sequence 48677, Application US/10831286A  
Publication No. US20060046246A1  
GENERAL INFORMATION:  
APPLICANT: ZENG, QIANDONG  
APPLICANT: CHATELIER, SONIA  
APPLICANT: MOIR, DONALD T.  
APPLICANT: LACROIX, BRUNA  
APPLICANT: CHILDESS, DARRELL  
TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
FILE REFERENCE: 032796-174.001  
CURRENT APPLICATION NUMBER: US/10/831.286A  
PRIOR FILING DATE: 2004-04-26  
PRIOR APPLICATION NUMBER: 60/464.955  
NUMBER OF SEQ ID NOS: 48788  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 48677  
LENGTH: 1507  
TYPE: DNA  
ORGANISM: *Providencia rettgeri*  
US-10-831-286A-48677

Query Match 69.9%; Score 1018; DB 11; Length 1507;  
Best Local Similarity 83.2%; Pred. No. 5,36-306;  
Matches 1217; Conservative 0; Mismatches 240; Indels 5; Gaps 5;

QY 1 ATGGAACGCTGGCGGAGATGCTTTACATGCAAGTCGAAACGGGAGAC-GGATGCTTGA 59  
 DB 11 ATGGAACGCTGGCGGAGATGCTTTACATGCAAGTCGAAACGGGAGACGTTGCT 70  
 QY 60 TCTGG-TGGCGAGTGGCGGAGCGGGTGAATGTCATCGAACGTATCCAGAAAGGGGG 118  
 DB 71 TCTCGCTGAGAGCGGCGGAGCGGGTGAATGTTATGGGGATCTGCCCGATAGAGGGGA 130  
 QY 119 TAAAGCATGAAAGATGTGCTAATCCGCAATATCTTAAAGAGGAAAGAGGGGATCGA 178  
 DB 131 TAAACACTGGAAACGGTGGCTAATACCGCATATCTTAAAGAGGAAAGAGGGGAACTT 190  
 QY 179 AAGACCTTGCGCTTTTGGAGCGGCGGATGTCATTAAGTATGAGTGGGCTTAAAGGCTT 238  
 DB 191 CGGTCTTGCGCTTATGAGATGAACCAATATGGATTAAGTATGAGTGGGCTTAAAGGCTT 250  
 QY 239 ACCAAGGCGAGCATAGTATGATGCTGAGAGGAGCAGCAGCAGCTGGGACTGAGACA 298  
 DB 251 ACTTACGCGAGCATCTCCAGTGTGTCTGAGAGGATGATCAGCAGCTGGGACTGAGACA 310  
 QY 299 CGGCGCAGACTCTCTACCGGAGGAGCAGCATGGGGAATTTTGGACAAATGGCGCAAGCTGA 358  
 DB 311 CGGCGCAGACTCTCTACCGGAGGAGCAGCATGGGGAATTTTGGACAAATGGCGCAAGCTGA 370  
 QY 359 TCCAGCAATGGCGCGTGAAGTGAAGAGGCGCTTGGGCTTAAAGCTTTCACTGAGCAA 418  
 DB 371 TCGAGCATGCGCGGTATGAAGAGGCGCTTGGGCTTAAAGCTTTCACTGAGCAA 430  
 QY 419 GAAAGGTTACGCTTAAATATCGTGAACCATGACGCTATCGAGAGAGAGACCGGCTT 478  
 DB 431 GAAAGGCTTCACTTAAATATCATCAACGATTAAGCTTACGAGAGAGAGAGACCGGCTT 490  
 QY 479 ACTACGTCCAGCAGACCGCGGCTAATACGATGGGTGCAAGGCTTAACTGGAATCTGGGC 538  
 DB 491 ACTCGGTCCAGCAGACCGCGGCTAATACGATGGGTGCAAGGCTTAACTGGAATCTGGGC 550  
 QY 539 GTTAAAGGTGGCGAGCGGCTTGTAAATGATGTAATCCCGGCGCTTAACTGGGA 598  
 DB 551 GTTAAAGGTGGCGAGCGGCTTGTAAATGATGTAATCCCGGCGCTTAACTGGGA 610  
 QY 599 ATTGCGTTGAAACTAACAAGCTAGAGTGTGGCAGAGGAGTGAATTCATGTGTAGC 658  
 DB 611 ATGGCATCTAAGCATGTGTGACGTAGAGTCTTGTAAAGGGGGGTGAATTCATGTGTAGC 670  
 QY 659 AGTGAATGTGTAGATATGAAGAAATCATGATGCGAAGGAGCGCTCTGGGCTTAA 718  
 DB 671 GTTGAATGTGTAGATATGAAGAAATCATGATGCGAAGGAGCGCGCTCTGGGCTTAA 730  
 QY 719 CTGACGCTCATGACCAAGACGTGGGAGGAGAAACAGGATTAAGTATCCCTGTGTAGCAG 778  
 DB 731 CTGACGCTCATGACGAGAAACGTGGGAGGAGAAACAGGATTAAGTATCCCTGTGTAGCAG 790  
 QY 779 CCCTAACGATGTC-AACTAGTGTGTGGGCTTATTAAGGCTTGTGTAAGAGCTTAAAGCGC 837  
 DB 791 CTGTAACGATGTCATTTGAGAGTGTGTCTTCTAGAGAGATGCGTTCCGAGAGCTTAAAGCGC 850  
 QY 838 TGAAGTTGACCGCTGGGAGTACGATCGCAAGTTAAACTCAAGGAAATGAAGCGGGA 897  
 DB 851 TTAATATGACCGCTGGGAGTACGATCGCAAGTTAAACTCAAGGAAATGAAGCGGGA 910  
 QY 898 CCGGCAACAGCGGTGATTAATGTGAGATTAATTCGCAACGCGAATAACCTTAACCTTACC 957  
 DB 911 CCGGCAACAGCGGTGATTAATGTGAGATTAATTCGCAACGCGAATAACCTTAACCTTACC 970  
 QY 958 CTTGACATGTAGCAATTTTCTAGAGATTAATTAAGTCTTGGGAAACGCTTAAACAGCT 1016  
 DB 971 CTTGACATGTAGCAATTTTCTAGAGATTAATTAAGTCTTGGGAAACGCTTAAACAGCT 1030  
 QY 1017 GCTGCATGAGTGTGTGAGCTGTGTGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGC 1076  
 DB 1031 GCTGCATGAGTGTGTGAGCTGTGTGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGC 1090  
 QY 1077 AACCTTGTATTAATGTCATC-ATTGTGTGGGCACTTAAATGAGACATGCGGCTGACA 1135

DB 1091 AACCTTATCTTTGTGTGCAAGCATACGATCGGGAACCTCAAGAGAGATGCGCGGTGATA 1150  
 QY 1136 AACCGAGGAAGAGTGGGGAATGACGTCAAGTCTCTAATGGCCCTTATGGGTAGGGCTTACACA 1195  
 DB 1151 AACCGAGGAAGAGTGGGGAATGACGTCAAGTCTCTAATGGCCCTTATGGGTAGGGCTTACACA 1210  
 QY 1196 CGTAAATACAAATGGCGGTACAGAGGTTGCCAACCAGCGGAGGGAGCTTAATCTCAGAAA 1255  
 DB 1211 CGTGTCTACAAATGGCGGTACAGAGGTTGCCAACCAGCGGAGGGAGCTTAATCTCAGAAA 1270  
 QY 1256 GCGCGTGTAGTCCGAGTGAAGTCTGCAATCTGCACTCGTGAATGCGAATGCTGATTA 1315  
 DB 1271 GTACGTGTAGTCCGAGTGAAGTCTGCAATCTGCACTCGTGAATGCGAATGCTGATTA 1330  
 QY 1316 ATGCGGATACAGATGTGCGGCTGAATACGTTCCCGGCTTGTACACACCGCGCTGAC 1375  
 DB 1331 ATGCGGATACAGATGTGCGGCTGAATACGTTCCCGGCTTGTACACACCGCGCTGAC 1390  
 QY 1376 ACCATGGAGTGGGTTTCAACAGAGCAGTATGCTTAACCTTAAGAGAGGCGCTTGGCAC 1435  
 DB 1391 ACCATGGAGTGGGTTTCAACAGAGCAGTATGCTTAACCTTACGAGAGGCGCTTACAC 1450  
 QY 1436 GGTGAGATTCACTGAGTGGGTG 1457  
 DB 1451 TTTGTGATTCACTGAGTGGGTG 1472

RESULT 11  
 US-10-831-286A-48684  
 ; Sequence 48684, Application US/10831286A  
 ; Publication No. US20060046246A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZENG, QIANDONG  
 ; APPLICANT: CHATELIER, SONIA  
 ; APPLICANT: MOIR, DONALD T.  
 ; APPLICANT: LACROIX, BRUNA  
 ; APPLICANT: CHILDRESS, DARELL  
 ; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
 ; FILE REFERENCE: 032796-174.001  
 ; CURRENT APPLICATION NUMBER: US/10/831,286A  
 ; PRIOR FILING DATE: 2004-04-26  
 ; PRIOR APPLICATION NUMBER: 60/464,955  
 ; PRIOR FILING DATE: 2003-04-24  
 ; NUMBER OF SEQ ID NOS: 48788  
 ; SOFTWARE: Patentin version 3.2  
 ; SEQ ID NO 48684  
 ; LENGTH: 1508  
 ; TYPE: DNA  
 ; ORGANISM: Providencia rustigianii  
 US-10-831-286A-48684

Query Match 69.8%; Score 1017; DB 11; Length 1508;  
 Best Local Similarity 83.2%; Pred. No. 1,1e-305;  
 Matches 1217; Conservative 0; Mismatches 240; Indels 6; Gaps 5;

QY 1 ATGGAACGCTGGCGGAGATGCTTTACATGCAAGTCGAAACGGGAGAC-GGATGCTTGA 59  
 DB 11 ATGGAACGCTGGCGGAGATGCTTTACATGCAAGTCGAAACGGGAGACGTTGCT 70  
 QY 60 TCTGG-TGGCGAGTGGCGGAGCGGGTGAATGTCATCGAACGTATCCAGAAAGGGGG 118  
 DB 71 TCTCGCTGAGAGCGGCGGAGCGGGTGAATGTTATGGGGATCTGCCCGATAGAGGGGA 130  
 QY 119 TAAAGCATGAAAGATGTGCTAATCCGCAATATCTTAAAGAGGAAAGAGGGGATCGA 178  
 DB 131 TAAACACTGGAAACGGTGGCTAATACCGCATATCTTAAAGAGGAAAGAGGGGAACTT 190  
 QY 179 AAGACCTTGCGCTTTTGGAGCGGCGGATGTCATTAAGTATGAGTGGGCTTAAAGGCTT 238  
 DB 191 CGGTCTTGCGCTTATGAGATGAACCAATATGGATTAAGTATGAGTGGGCTTAAAGGCTT 250  
 QY 239 ACCAAGGCGAGCATAGTATGATGCTGAGAGGAGCAGCAGCAGCTGGGACTGAGACA 298  
 DB 251 ACTTACGCGAGCATCTCCAGTGTGTCTGAGAGGATGATCAGCAGCTGGGACTGAGACA 310  
 QY 299 CGGCGCAGACTCTCTACCGGAGGAGCAGCATGGGGAATTTTGGACAAATGGCGCAAGCTGA 358  
 DB 311 CGGCGCAGACTCTCTACCGGAGGAGCAGCATGGGGAATTTTGGACAAATGGCGCAAGCTGA 370  
 QY 359 TCCAGCAATGGCGCGTGAAGTGAAGAGGCGCTTGGGCTTAAAGCTTTCACTGAGCAA 418  
 DB 371 TCGAGCATGCGCGGTATGAAGAGGCGCTTGGGCTTAAAGCTTTCACTGAGCAA 430  
 QY 419 GAAAGGTTACGCTTAAATATCGTGAACCATGACGCTATCGAGAGAGAGACCGGCTT 478  
 DB 431 GAAAGGCTTCACTTAAATATCATCAACGATTAAGCTTACGAGAGAGAGAGACCGGCTT 490  
 QY 479 ACTACGTCCAGCAGACCGCGGCTAATACGATGGGTGCAAGGCTTAACTGGAATCTGGGC 538  
 DB 491 ACTCGGTCCAGCAGACCGCGGCTAATACGATGGGTGCAAGGCTTAACTGGAATCTGGGC 550  
 QY 539 GTTAAAGGTGGCGAGCGGCTTGTAAATGATGTAATCCCGGCGCTTAACTGGGA 598  
 DB 551 GTTAAAGGTGGCGAGCGGCTTGTAAATGATGTAATCCCGGCGCTTAACTGGGA 610  
 QY 599 ATTGCGTTGAAACTAACAAGCTAGAGTGTGGCAGAGGAGTGAATTCATGTGTAGC 658  
 DB 611 ATGGCATCTAAGCATGTGTGACGTAGAGTCTTGTAAAGGGGGGTGAATTCATGTGTAGC 670  
 QY 659 AGTGAATGTGTAGATATGAAGAAATCATGATGCGAAGGAGCGCTCTGGGCTTAA 718  
 DB 671 GTTGAATGTGTAGATATGAAGAAATCATGATGCGAAGGAGCGCGCTCTGGGCTTAA 730  
 QY 719 CTGACGCTCATGACCAAGACGTGGGAGGAGAAACAGGATTAAGTATCCCTGTGTAGCAG 778  
 DB 731 CTGACGCTCATGACGAGAAACGTGGGAGGAGAAACAGGATTAAGTATCCCTGTGTAGCAG 790  
 QY 779 CCCTAACGATGTC-AACTAGTGTGTGGGCTTATTAAGGCTTGTGTAAGAGCTTAAAGCGC 837  
 DB 791 CTGTAACGATGTCATTTGAGAGTGTGTCTTCTAGAGAGATGCGTTCCGAGAGCTTAAAGCGC 850  
 QY 838 TGAAGTTGACCGCTGGGAGTACGATCGCAAGTTAAACTCAAGGAAATGAAGCGGGA 897  
 DB 851 TTAATATGACCGCTGGGAGTACGATCGCAAGTTAAACTCAAGGAAATGAAGCGGGA 910  
 QY 898 CCGGCAACAGCGGTGATTAATGTGAGATTAATTCGCAACGCGAATAACCTTAACCTTACC 957  
 DB 911 CCGGCAACAGCGGTGATTAATGTGAGATTAATTCGCAACGCGAATAACCTTAACCTTACC 970  
 QY 958 CTTGACATGTAGCAATTTTCTAGAGATTAATTAAGTCTTGGGAAACGCTTAAACAGCT 1016  
 DB 971 CTTGACATGTAGCAATTTTCTAGAGATTAATTAAGTCTTGGGAAACGCTTAAACAGCT 1030  
 QY 1017 GCTGCATGAGTGTGTGAGCTGTGTGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGC 1076  
 DB 1031 GCTGCATGAGTGTGTGAGCTGTGTGTGAGATGTTGGGTTAAAGTCCCGCAACGAGCGC 1090  
 QY 1077 AACCTTGTATTAATGTCATC-ATTGTGTGGGCACTTAAATGAGACATGCGGCTGACA 1135

Db 251 ACCAAGCGACGATCCCTTACTGTGTGAGAGATGATCAGCACACTGGGACTGAGACA 310  
Qy 299 CGGCCCCAGACTCTTACGAGAGGAGCAGATGGGAAATTTTGACATATGGCGCAAGCCTGA 358  
Db 311 CGGCCCCAGACTCTTACGAGAGGAGCAGATGGGAAATTTTGACATATGGCGCAAGCCTGA 370  
Qy 359 TCCAGCAATGCCGCTGATGAGAGAGCCTTGGGGTTGTAAGCTTTTCAGTCCAGAA 418  
Db 371 TGCAGCATGCCGCTGATGAGAGAGCCTTGGGGTTGTAAGCTTTTCAGTGGAG 430  
Qy 419 GAAAGGTTACGGTAAATTAATCGGACCATGACCGGTATGACAGAGAGACCGGCTA 478  
Db 431 GAAAGGCTTATGCTATATATCATCATGAGATTGACGTTACCAAGAGAGACCGGCTA 490  
Qy 479 ACTACGTCCAGAGACCGCGGTAAATCATGAGGTGCAAGCGTTAATCGAATTAATCTGGC 538  
Db 491 ACTCCGTGCAAGACCGCGGTAAATCATGAGGTGCAAGCGTTAATCGAATTAATCTGGC 550  
Qy 539 GTAAAGGTGCGCAGCGCGCTTGTAAATGTCATGATGTGAATCCCGGGCTTAACCTGGGA 598  
Db 551 GTAAAGGTGCGCAGCGCGCTTGTAAATGTCATGATGTGAATCCCGGGCTTAACCTGGGA 610  
Qy 599 ATTGCGTTTGAATCTAAGAGCTAGAGTGTGCGAGAGGAGTGGAAATTCATGATGAGC 658  
Db 611 ATGACATCTAAGACTGTGCTAGCTAGACTCTTGTAGAGAGGAGTGAATTCATGATGAGC 670  
Qy 659 AGTGAATGCGTATGAGATATGAGAGAAATCATGATGCGAAGCAGCCTCTGGGTAAACA 718  
Db 671 GGTGAATGCGTATGAGATATGAGAGAAATCATGATGCGAAGCAGCCTCTGGGTAAACA 730  
Qy 719 CTGACGCTCATGACGAGAGCGTGGGAGCAAAAGAGATTAGATACCTGGTATGCCAG 778  
Db 731 CTGACGCTCATGAGTGCAGAAAGCGTGGGAGCAAAAGAGATTAGATACCTGGTATGCCAG 790  
Qy 779 CCTTAAACGATGTCACTAGTGTGTGGCCTTATTAAGCTTGG-TAAAGAGCTTAAGCGC 837  
Db 791 CTGTAAACGATGTCACTAGTGTGTGGCCTTATTAAGCTTGG-TAAAGAGCTTAAGCGC 850  
Qy 838 TGAAGTTGACCGCGCTGGGAGTACGTCGCAAGATTAAACTCAAGAGAAATTTGACGGGGA 897  
Db 851 TTTAAATGACCGCGCTGGGAGTACGTCGCAAGATTAAACTCAAGAGAAATTTGACGGGGA 910  
Qy 898 CCGGCAAGCGGTGATTTATGTGAGTTAAATTCGATCAACGCGAAAACTTAACCTTACC 957  
Db 911 CCGGCAAGCGGTGATTTATGTGAGTTAAATTCGATCAACGCGAAAACTTAACCTTACC 970  
Qy 958 CTGACATGTAGCGCAATTTTCTAGAGTATGATTTAGT-CTTCGGGAAAGCTTAACAGGT 1016  
Db 971 CTGACATGTAGCGCAATTTTCTAGAGTATGATTTAGTCTTTCGGGAACTCTGAGACAGGT 1030  
Qy 1017 GGTGACATGCTGTGTGAGCTCGTGTGAGTATGATTTAGTGGGTAAAGTCCCGCAAGAGCGC 1076  
Db 1031 GGTGACATGCTGTGTGAGCTCGTGTGAGTATGATTTAGTGGGTAAAGTCCCGCAAGAGCGC 1090  
Qy 1077 AACCTTGTATTAATTTGCAATCA--TTTGGTTGGGCACTTTAATGAGACTGCCGTGAC 1134  
Db 1091 AACCTTATCTTTTGTGTGCAAGCTCATGATGTGGAACTCAAGAGAACTGCGCGGTGAT 1150  
Qy 1135 AAACCGGAGAGAGGTGGGAGTATGATGATGATTTTCTATGAGCTTATGAGGCTTAC 1194  
Db 1151 AAACCGGAGAGAGGTGGGAGTATGATGATGATTTTCTATGAGCTTATGAGGCTTAC 1210  
Qy 1195 ACGTAATACATGCGGTGATGAGAGGTGTCACACCGCGAGGGGAGGCTAATCTCAGAA 1254  
Db 1211 ACGGTACATACGCGGTGATGAGAGGTGTCACACCGCGAGGGGAGGCTAATCTCAGAA 1270  
Qy 1255 AGCGCGTGTAGTCCGATCCGAGTCTGCACTCGACTCCGTGAGTCCGATCCGTAGT 1314  
Db 1271 AGTACGTGTAGTCCGATCCGAGTCTGCACTCGACTCCGTGAGTCCGATCCGTAGT 1330  
Qy 1315 AATCGGATCAGATGTCCGGTGATTAAGTCTTCCGGGTCTTGTACACACCGCGCTCA 1374

Db 1331 AATGTAGATACAGATGCTACGCTGATATGTTCCCGGGCCTTGTACACACCGCGCTCA 1390  
Qy 1375 CACCATGGAGAGGTGGTTTACCAAGAGCAGGTAGTCTAAACCGTAAGAGGCGCTTGCA 1434  
Db 1391 CACCATGGAGAGGTGGTTTACCAAGAGAGTATGTTTAACTTTCGGAGGCGCTTACCA 1450  
Qy 1435 CGGTGAGATTATCATGCTGGGGTG 1457  
Db 1451 CTTTGTGATTCATGACTGGGGTG 1473

RESULT 12  
US-10-831-286A-48682  
; Sequence 48682, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILHRES, DARRELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; CURRENT FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; PRIOR FILING DATE: 2003-04-24  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 48682  
; LENGTH: 1510  
; TYPE: DNA  
; ORGANISM: *Mesliereella wisconsinensis*  
; US-10-831-286A-48682

Query Match 69.7%; Score 1015.4; DB 11; Length 1510;  
Best Local Similarity 82.7%; Pred. No. 3,4e-305;  
Matches 1211; Conservative 0; Mismatches 246; Indels 8; Gaps 4;

Qy 1 ATGGAAGCTGTGGCGGACTGCTTATACATGACATGCGAAGCGAGCAGCAGATGCTTGCA- 59  
Db 11 ATGGAAGCTGTGGCGGAGGCTTAAACATGACATGCGAAGCGAGGATGCTTGC 70  
Qy 60 ---TTGTGTGCGAGATGGCGGAGCGGTGAGTATGATCGGAACGATCCAGAAAGAGGG 116  
Db 71 TTTCTTGTGTCAGCAGAGCGGCGGAGCGGTGAGTATGATGAGGATCTGCTGACAGAGGG 130  
Qy 117 GGTAAAGCATCGAAGATGTGCTAATACGSCATATCTTAAGAGAGAAAGCAGGGGATC 176  
Db 131 GATTAACATACGGAAGACGATGATTAATACGSCATATCTTAAAGAGCAAGAGGGAGCC 190  
Qy 177 GAAAGACCTTGCCCTTTTGGAGCGGCGGATGTCTGATTTAGCTAGTTGTGGGTAAAGGC 236  
Db 191 TTTGGGCTTGGCGCTGTGCGATGAACCCATATGGATTTAGTATGATGATGATGATG 250  
Qy 237 CTACCAAGGAGGAGATCAATGATGTTGTCTGAGAGAGCAGCAGCCACACTGGGACTGAGA 296  
Db 251 TCACTTAAGGAGGAGATCTAGCTGATGAGAGATGATCAAGCAGCAGCTGGGACTGAGA 310  
Qy 297 CAGGCGCCAGACTCTTAAGGAGGAGCAGCATTTGGGAAATTTTGGACATATGGGCGCAAGCT 356  
Db 311 CAGGCGCCAGACTCTTAAGGAGGAGCAGCATTTGGGAAATTTTGGACATATGGGCGCAAGCT 370  
Qy 357 GATCCAGCAATGCGCGCTGATGAGAGAGGCTTTCGGGTGTAAAGCTCTTTCAGTCCAG 416  
Db 371 GATCCAGCAATGCGCGCTGATGAGAGAGGCTTTCGGGTGTAAAGCTCTTTCAGTCCAG 430  
Qy 417 AAGAAAGGTTTACGTTAATATCTGATGAGCCATGACGCTATCGACAGAGAGAGCAGCCGC 476  
Db 431 AGAAGGCGTTGATTAATTAATATCATGACGATTAAGCTTACCGACAGAAAGCAGCCGC 490  
Qy 477 TAACTAGTTCAGCAGAGCCCGGTAATACCTAGAGTCCAGCGCTTAATCGGAATTAAGT 536







```
QY 777 CCCTTAAACGATGTCACACTAGTGTGTTGGGCTTATTAGCTTGG- TAAAGAACTAACG 835
Db 802 CGCCGTAACGATGTCGACTTGGAGGTTGTGCCCTTGAAGCGGTGCTTCCGAGGCTTAACG 861
QY 836 CGTGAAGTTGACCGCTCGGGGAGTAAAGTTCGCAAGTTAAATCTCAAGAAATTGACGGG 895
Db 862 CTTTAAGTCGACCGCTCGGGGAGTAAAGTTCGCAAGTTAAATCTCAAGAAATTGACGGG 921
QY 896 GACCGGCAAGGGGAGTAAATGATTAATTGATGCAAGCAAGCAAGAAACCTTACTTA 955
Db 922 GGCCTGACAGAGGGTGAAGCATGTGTTTAATTGATGCAAGCAAGCAAGAACTTACTTA 981
QY 956 CCCTTGAATGATGCAAGTATTTCTAGAGATAGATTAATG- CTTGCGGAAAGCTTAACACAG 1014
Db 982 CTTTGAATGATGCAAGTATTTCTAGAGATAGATTAATG- CTTGCGGAAAGCTTGAAGACAG 1041
QY 1015 GTGCTGATGCTGTGCTGACGCTGTGTGTTGATGATGTTGGTTAAAGTCCCGAACAGAC 1074
Db 1042 GTGCTGATGCTGTGCTGACGCTGTGTGTTGATGATGTTGGTTAAAGTCCCGAACAGAC 1101
QY 1075 GCAACCTTGTCAATTAATGCAATC- ATTTGTTGGGCACTTAAATGACCTGCGGTGA 1133
Db 1102 GCAACCTTATCTTTGTTGGCAAGCGGTCCGCGGAACTCAAGAGACCTGCACTAT 1161
QY 1134 CAACCGAGAGAGTGGGATGACGTCAGTCAAGTCTCATGCGCTTATGAGGTAGGGCTTCA 1193
Db 1162 TAAATCTGAGAGAGTGGGATGACGTCAGTCAAGTCTCATGCGCTTATGAGGTAGGGCTTCA 1221
QY 1194 CACGTAATCAATGCGCGGTACAGAGGGTTGCCAACCCCGAGAGGGAGCTTAATCTCAGA 1253
Db 1222 CACGTCATCAATGCGCGGTACAGAGGGTTGCCAACCCCGAGAGGGAGCTTATCTCAGA 1281
QY 1254 AAGCGGTGTAGTCCGAGTCCGAGTCTGCAATCTCCGAGAGTCCGAGTCCGAGTCCGAGT 1313
Db 1282 AAGTGTGTGTAGTCCGAGTCCGAGTCTGCAATCTCCGAGAGTCCGAGTCCGAGTCCGAGT 1341
QY 1314 TAAATGCGAGTCAAGTGTGCGGTGAATACGTTCCCGGGCTTGTATACACCGCCGCTC 1373
Db 1342 TAAATGCGAGTCAAGTGTGCGGTGAATACGTTCCCGGGCTTGTATACACCGCCGCTC 1401
QY 1374 ACACATGAGAGTGGGTTTTCACAGAGAGAGTATGTTAAACGTTAAGAGAGGCGCTTGC 1433
Db 1402 ACACATGAGAGTGGGTTTTCACAGAGAGAGTATGTTAAACGTTAAGAGAGGCGCTTACC 1461
QY 1434 ACCGTGATTCATGACTGGGGTG 1457
Db 1462 ACTTGTGATTCATGACTGGGGTG 1485

RESULT 14
US-10-831-286A-48686
; Sequence 48686, Application US/10831286A
; Publication No. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48686
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Yokenella regensburgeri
US-10-831-286A-48686
```

```
Query Match 69.1%; Score 1007.2; DB 11; Length 1492;
Best Local Similarity 82.4%; Pred. No. 12e-302; Indels 5; Gaps 4;
Matches 1204; Conservative 0; Mismatches 253;

QY 1 ATTGAACGCTGCGGCATGCTTTTACATGCAATGCAACGCGACGACGATGCTTGCAT 60
Db 1 ATTGAACGCTGCGGCATGCTTTTACATGCAATGCAACGCGACGACGATGCTTGCAT 60
QY 61 CT -GGTGGAGTGGCGGACGCGGTGATGATGCAATCGCAACGTATCCAGAAAGAGGGGG 118
Db 61 CTCGGGTGACAGTGGCGGACGCGGTGATGATGCAATCGCAACGCGGTGATGCGGGGA 120
QY 119 TAAAGCATGAAAGATGCTTAATACCGCATATACCTTAAGAGAGAAAGCAGGGATGCA 178
Db 121 TAAAGCATGAAAGATGCTTAATACCGCATATACCTTAAGAGAGAAAGCAGGGATGCA 180
QY 179 AAGACCTTGGCGCTTTTGGAGCGGCGATGTCTGATTAAGCTAGTGGGTAAAGGCT 238
Db 181 CGGCGCTTCTTGCATCAGATGTCGCCAGATGGGATTAAGTATGAGGTAAAGGCTC 240
QY 239 ACCAAGGCGACGATCAGTATGTTGTTGAGAGACGACCAAGCTGAGTGAACA 298
Db 241 ACTTAAGGCGACGATCCTTGTGCTGTGAGAGATGACCAAGCTGAGTGAACA 300
QY 299 CGGCGCAAGCTCTTACGAGAGGACGACGATGGGGAATTTTGAACAATGGCGCAAGCTGA 358
Db 301 CGGTCGCAAGCTCTTACGAGAGGACGACGATGGGGAATTTTGAACAATGGCGCAAGCTGA 360
QY 359 TCCAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTGTTAAAGCTCTTCACTGCAAA 418
Db 361 TCGAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTGTTAAAGCTCTTCACTGCAAA 420
QY 419 GAAAGGTATACGGTAAATATGCTGACCCATGACGCTATGACAGAAAGACACGCGCTA 478
Db 421 GAAAGGTATACGGTAAATATGCTGACCCATGACGCTATGACAGAAAGACACGCGCTA 480
QY 479 ACTACGTGCAAGCAGCGCGGTATATGTAAGGTGCAAGCGTATGACAGAAAGACACGCGCTA 538
Db 481 ACTACGTGCAAGCAGCGCGGTATATGTAAGGTGCAAGCGTATGACAGAAAGACACGCGCTA 540
QY 539 GTTAAAGGTGCGAGCGCGCTTGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 598
Db 541 GTTAAAGGTGCGAGCGCGCTTGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 600
QY 599 ATTGCGTTGAAACTCAAAAGCTAGAGTGTGAGAGAGAGAGTGAATTCATGTTAGC 658
Db 601 ACTGCAATTCGAAATCTGACAGGCTTGAATGCTTGTAAAGAGAGAGAGTGAATTCATGTTAGC 660
QY 659 AGTGAATGCTGAGATATGAAAGCAATCGATGCGAAAGGACGCTCTGAGTTAAACA 718
Db 661 GTTGAATGCTGAGATATGAAAGCAATCGATGCGAAAGGACGCTCTGAGTTAAAGCA 720
QY 719 CTGACGCTCATGACAGAAAGCTGGGAGAGCAACAGATTAATGTAATGTAATGTAATGTAATGTAAT 778
Db 721 CTGACGCTCATGAGTGCAGAAAGCTGGGAGAGCAACAGATTAATGTAATGTAATGTAATGTAATGTAAT 780
QY 779 CCCTTAAAGCATGTCATGATGTTGTTGGGCTTATTAAGGCTTGG- TAAAGAGCTTAACGCG 837
Db 781 CCCTTAAAGCATGTCATGATGTTGTTGGGCTTATTAAGGCTTGG- TAAAGAGCTTAACGCG 840
QY 838 TGAAGTTGACCGCTCGGGAGTACGCTCGCAAGATTAACCTCAAGAAATTGACGGGGA 897
Db 841 TTAAGTTGACCGCTCGGGAGTACGCTCGCAAGATTAACCTCAAGAAATTGACGGGGA 900
QY 898 CCGGACACAGCGGTGATTAATGATGATTAATGATGCAACGCGGAAACCTTAACCTAC 957
Db 901 CCGGACACAGCGGTGATTAATGATGATTAATGATGCAACGCGGAAACCTTAACCTAC 960
QY 958 CTTGACATGTAAGGAATTTCTAGAGTATGATTAAGT- CTTGCGGAAAGCTTAACAGAGT 1016
Db 961 CTTGACATGTAAGGAATTTCTAGAGTATGATTAAGT- CTTGCGGAAAGCTTAACAGAGT 1020
```

QY 1017 GCTGATGCTGTGCTCACTGCTGCTGAGATGTTGGGTTAAGTCCCGCAAGAGGCG 1076  
DB 1021 GCTGATGCTGTGCTCACTGCTGCTGAGATGTTGGGTTAAGTCCCGCAAGAGGCG 1080  
QY 1077 AACCTTGTCAATTAATGCGCATC-ATTGGTTGGGCACTTTAATGAGACTGCGGTGACA 1135  
DB 1081 AACCTTATCTTTGTGTGCGACGGGTCCGCGGGAACCTCAAGAGAGACTGCGAGTAA 1140  
QY 1136 AACCGAGAGAGTGGGAGTGAAGTCAAGTCTCATGCGCTTTATGGGTGGCTTCA 1195  
DB 1141 AACTGAGAGAGTGGGAGTGAAGTCAAGTCTCATGCGCTTTATGGGTGGCTTCA 1200  
QY 1196 CGTAATACATGCGCGCTACAGAGGTTGCCACCCCGAGAGGGAGAGCTAATCTAGAA 1255  
DB 1201 CGTGTCAATGCGCATTAACAAGAGAGAGAGCTTCGAGAGAGAGAGAGCTTCA 1260  
QY 1256 GCGCGTGTGATCGCGGATCGAGTCTGCAACTGCACTCCGTAAGTCGAGATCGTAGTA 1315  
DB 1261 GATGTGTGATCGCGGATCGAGTCTGCAACTGCACTCCGTAAGTCGAGATCGTAGTA 1320  
QY 1316 ATGCGGATACAGCATGTGCGGTGAATACGTTCCCGGCTCTTTACACACCGCGCTCAC 1375  
DB 1321 ATGCGGATACAGCATGTGCGGTGAATACGTTCCCGGCTCTTTACACACCGCGCTCAC 1380  
QY 1376 ACCATGGAGTGGGTTTACACAGAGAGTGTCTAACCCTAAGAGGAGGCGCTTCCAC 1435  
DB 1381 ACCATGGAGTGGGTTTACACAGAGAGTGTCTAACCCTAAGAGGAGGCGCTTACAC 1440  
QY 1436 GGTGAGATTCAATGATCGGGTG 1457  
DB 1441 TTTGTGATTCAATGATCGGGTG 1462

RESULT 15  
US-10-831-286A-48687  
; Sequence 48687, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILDRESS, DARRELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; PRIOR FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 48687  
; LENGTH: 1519  
; TYPE: DNA  
; ORGANISM: Yokenella regensburgei  
US-10-831-286A-48687

Query Match 69.1%; Score 1007.2; DB 11; Length 1519;  
Best Local Similarity 82.4%; Pred. No. 1.2e-302;  
Matches 1204; Conservative 0; Mismatches 253; Indels 5; Gaps 4;  
QY 1 ATTGAACGCTGGCGGATCTTTACATGCAAGTGAAGCGAGACGAGATGCTTGCA 60  
DB 22 ATTGAACGCTGGCGGATCTTTACATGCAAGTGAAGCGAGACGAGATGCTTGCT 81  
QY 61 CT--GGTGGCGAGTGGCGGAGTGAATGCAATCGGAAGTATCCAGAGAGAGGGGG 118  
DB 82 CTCGGGTGAGAGTGGCGGAGTGAATGCTGAGAACTGCTGATGAGAGGGGA 141  
QY 119 TAAACCATGAAAGATGTCTAATACCGCATATCTCTAAGAGAGAGAGAGAGATCGA 178  
DB 142 TAACTACTGAAACGGTAGCTAATACCGCATATCGTGCAGAGACCAAGAGGGGACCTT 201

QY 179 AAGACCTTCGCTTTTGGAGCGGCGATGTCATTAAGTGGTGGGTTAAGGCGCT 238  
DB 202 CGGGCTCTTGGCCATCAGATGTGCCAGATGGAGTTAGTGTGGGTTAAGGCGCT 261  
QY 239 ACCAAGGAGAGATCAATGATTTGTGAGAGAGACGACCGACACTGGAGCTGAGACA 298  
DB 262 ACTTGGGAGAGATCCCTAGCTGTGAGAGAGTGAACAGACCACTGGAATCTGAGACA 321  
QY 299 CGGCCCAATCTCTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358  
DB 322 CGGTCAACATCTCTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 381  
QY 359 TCCAGCAATGCGCGGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418  
DB 382 TCCAGCAATGCGCGGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 441  
QY 419 GAAAGGTTACGATTAATATGTTGATCCCATGACGATGACAGAGAGAGAGAGAGAG 478  
DB 442 GAAAGGTTACGATTAATATGTTGATCCCATGACGATGACAGAGAGAGAGAGAGAG 501  
QY 479 ACTAGTCCAG 538  
DB 502 ACTAGTCCAG 561  
QY 539 GTAAAGGTTGCGAG 598  
DB 562 GTAAAGGTTGCGAG 621  
QY 599 ATTGGCTTGAATCTAAG 658  
DB 622 ACTGATTCGAATCTGAG 681  
QY 659 AGTGAATGCGTGAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718  
DB 682 GGTGAATGCGTGAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 741  
QY 719 CTGACGCTCATGACGAAAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778  
DB 742 CTGACGCTCATGACGAAAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 801  
QY 779 CCTTAAGATGTCATTAATGTTGGGCTTTATGAGCTTGG-TAAGAGAGTAAAGCG 837  
DB 802 CCTTAAGATGTCATTAATGTTGGGCTTTATGAGCTTGG-TAAGAGAGTAAAGCG 861  
QY 838 TGAATGTAAGCGCTGGGAGAGTACGATGCAAGATTAAGAGAGAGAGAGAGAGAGAG 897  
DB 862 TTAAGTGAAGCGCTGGGAGAGTACGATGCAAGATTAAGAGAGAGAGAGAGAGAGAG 921  
QY 898 CCGGCAAGAGGAGTGAATTTATGAGATTAATTCAGATGCAAGAGAGAGAGAGAGAGAG 957  
DB 922 CCGGCAAGAGGAGTGAATTTATGAGATTAATTCAGATGCAAGAGAGAGAGAGAGAGAG 981  
QY 958 CTGACATGAGAGATTTTCTAGAGATTAATTCAGATGCAAGAGAGAGAGAGAGAGAGAG 1016  
DB 982 CTGACATGAGAGATTTTCTAGAGATTAATTCAGATGCAAGAGAGAGAGAGAGAGAGAG 1041  
QY 1017 GCTGATGCTGTGCTCACTGCTGCTGCTGAGATTTGGGTTAAGTCCCGCAAGAGAGCG 1076  
DB 1042 GCTGATGCTGTGCTCACTGCTGCTGCTGAGATTTGGGTTAAGTCCCGCAAGAGAGCG 1101  
QY 1077 AACCTTGTCAATTAATGCGCATC-ATTGGTTGGGCACTTTAATGAGACTGCGGTGACA 1135  
DB 1102 AACCTTATCTTTGTGTGCGACGGTTCGCGGGAACCTCAAGAGAGAGAGAGAGAGAG 1161  
QY 1136 AACCGAGAGAGTGGGAGTGAAGTCAAGTCTCAAGAGAGAGAGAGAGAGAGAGAGAG 1195  
DB 1162 AACTGAGAGAGTGGGAGTGAAGTCAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAG 1221  
QY 1196 CGTAATACATGCGCGCTACAGAGGTTGCCACCCCGAGAGGGAGAGCTAATCTAGAA 1255  
DB 1222 CGTGTCAATGCGCATTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1281  
QY 1256 GCGCGTGTGATCGCGGATCGAGTCTGCAACTGCACTCCGTAAGTCCGTAAGT 1315



```

US-10-831-286A-48669
; Sequence 48669, Application US/10831286A
; Publication No. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48669
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Cedecea neteri
US-10-831-286A-48669

Query Match      69.0%; Score 1005.2; DB 11; Length 1494;
Best Local Similarity 82.2%; Pred. No. 5.1e-302;
Matches 1204; Conservative 0; Mismatches 253; Indels 7; Gaps 4;

QY 1 ATGAAAGCTGGCGGCTTATACATGCAAGTCGAAAGCGGACGAGCTTG---57
DB 1 ATTGAACGCTGGCGGCGCTTAAACATGCAAGTCGAAAGCGGACGAGCTTGCG 60
QY 58 -CATCTGGTGGCGAGTGGCGGAGCGGCTGAGTAATGCATCGAAAGTAACGAAGAGGG 116
DB 61 TCCTGGGTGACGAGCGGCGGAGCGGCTGAGTAATGTCGCGGAGATCTGCTGATGAGGG 120
QY 117 GGTAAAGCATCGAAAGATGCTTAATCCGCTAATACCTAAGAGAGAAAGAGGGGATC 176
DB 121 GATTAATCTAGTGAACCGTATGATTAATCCGCTAATACCTGCAAGAACCAAGAGGGGATC 180
QY 122 GATTAATCTAGTGAACCGTATGATTAATCCGCTAATACCTGCAAGAACCAAGAGGGGATC 180
DB 121 GATTAATCTAGTGAACCGTATGATTAATCCGCTAATACCTGCAAGAACCAAGAGGGGATC 180
QY 177 GATAAGCCTTGGCGCTTATGAGCGGCGGAGTGTCTGATTAATGCTAGTGGTGAAGAGC 236
DB 181 TTGGGGCTCTTGGCATCAATGATGAACCAAGATGGAGTTAGCTAGTGAAGTATGCG 240
QY 237 CTACCAAGCGGAGCATCAATGATGTTGTTCTGAGAGAGCAGCAGCAGCATCTGGAGCTGAGA 296
DB 241 TCACCTAGGAGGAGATCCTAGCTGCTGAGAGAGTGAACAGCAGCATCTGGAGCTGAGA 300
QY 297 CAGGCGCCAGATCTCTACCGGAGGCGAGCATGGGGATTTTGGACATATGGCGGAGAGCT 356
DB 301 CAGGCTCCAGATCTCTACCGGAGGCGAGCATGGGGATTTTGGACATATGGCGGAGAGCT 360
QY 357 GATCCAGCAATGCGCGCTGAGTGAAGAGCCTTGGGTTGTAAGCTCTTTTCAGTGCAG 416
DB 361 GATCCAGCAATGCGCGCTGAGTGAAGAGCCTTGGGTTGTAAGCTCTTTTCAGTGCAG 420
QY 417 AAGAAAAGGTTAAGCGTAATATCTGAGACCATGACGGTATCGACAGAAAGACCGGCG 476
DB 421 AAGAAAAGGTTAAGCGTAATATCTGAGACCATGAGCTTACTCGCAAGAAAGACCGGCG 480
QY 477 TAACTAGTGGCGAGCAGCGCGGTTAAAGTGAAGGTTGCAAGCGTTATCGGAATTAATG 536
DB 481 TAACTAGTGGCGAGCAGCGCGGTTAAAGTGAAGGTTGCAAGCGTTATCGGAATTAATG 540
QY 537 GCGTAAAGGGTGGCGAGCGGCGCTTGTAACTGAGATGGAATATCCCGGCTTAACCTGG 596
DB 541 GCGTAAAGGGTGGCGAGCGGCGCTTGTAACTGAGATGGAATATCCCGGCTTAACCTGG 600
QY 597 GAATGGCTTTGAAGTAACTGAGATGAGTGGCAGAGGAGTGAATTCATGTGTA 656
DB 601 GAATGGCTTTGAAGTAACTGAGATGAGTGGCAGAGGAGTGAATTCATGTGTA 660
QY 657 GCAATGAATGCGTGAAGATGAGTGAAGAAATCATGATGCGGAAGCAGCTCTCTGGTTAA 716

```

```

DB 661 GCGGTGAATATGCGTGAAGATCTGAGGAATACCGGTGGCGAAGCGGCCCTCGACAAA 720
QY 717 CACTGACGCTCATGCAACGAAGCGTGGGAGAGCAAGATTAATGATACCTGTAGTCA 776
DB 721 GACTGACGCTCATGCAAGCGTGGGAGAGCAAGATTAATGATACCTGTAGTCA 780
QY 777 GCGCCCTAAAGCATGTCATTAAGTGTGGGCTTTATAGGCTTGG-TAAGAACTAGC 835
DB 781 CCGCTTAAGCATGTCATTAAGTGTGGGCTTTATAGGCTTGG-TAAGAACTAGC 840
QY 836 CGTGAAGTGAACCGCTGGGAGTACGCTGCAAGATTAATCTAAGAAATGACGG 895
DB 841 CGTTAAGTGAACCGCTGGGAGTACGCTGCAAGATTAATCTAAGAAATGACGG 900
QY 896 GACCGGCAACAGCGGTGATTAATGAGTGAATTAATGATGCAACCGAAACCTTACCTA 955
DB 901 GCGCGGCAACAGCGGTGATTAATGAGTGAATTAATGATGCAACCGAAACCTTACCTA 960
QY 956 CCCTGACATGTAAGGGAATTTTCTAGATAGATTAAGTGTG-TTCGGGAACGCTAACACAG 1014
DB 961 CTCTGACATGTAAGGGAATTTTCTAGATAGATTAAGTGTG-TTCGGGAACGCTAACACAG 1020
QY 1015 GTGCTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074
DB 1021 GTGCTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1075 GCAACCTTGTCTATTAATGCTATC-ATTGGTGGGCACTTTAATGACATCGCGCTGA 1133
DB 1081 GCAACCTTGTCTATTAATGCTATC-ATTGGTGGGCACTTTAATGACATCGCGCTGA 1140
QY 1134 CAAACCGGAGGAAGGTTGGGAGTGAACGTAAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1193
DB 1141 TAACTGAGGAAGGTTGGGAGTGAACGTAAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1194 CAGTAAATCAATGCGCGCTGACAGAGGTTGCCAACCGCGAGGAGACTAATCTCAGA 1253
DB 1201 CAGTAAATCAATGCGCGCTGACAGAGGTTGCCAACCGCGAGGAGACTAATCTCAGA 1260
QY 1254 AAGCGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1313
DB 1261 AAGCGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1314 TAACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1373
DB 1321 TAACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1380
QY 1374 ACACATGGAAGTGGGTTTACCAAGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1433
DB 1381 ACACATGGAAGTGGGTTTACCAAGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1440
QY 1434 ACGGTGAATTCATGACTGGGCTG 1457
DB 1441 ACTTGTGATTCATGACTGGGCTG 1464

RESULT 18
US-10-831-286A-48665
; Sequence 48665, Application US/10831286A
; Publication No. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: PatentIn version 3.2

```

```
; SEQ ID NO 48665
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Cedecia davisiae
US-10-831-286A-48665

Query Match      69.0%; Score 1005.2; DB 11; Length 1503;
Best Local Similarity 82.2%; Pred. No. 5.1e-302;
Matches 1204; Conservative 0; Mismatches 253; Indels 7; Gaps 4;

QY 1 ATTGAACGCTGGGCGCATGCTTTACATGCAAGTGAAGGCGACGCGATGCTTGA- 59
DB 2 ATTGAACGCTGGGCGCGCCTTAACATGCAAGTGAAGGCGATGCAAGGATGCTTGC 61
QY 60 ---TCTGGTGGCGAGTGGCGGACGCGGTGAATGATCCGAAAGTATCCAGAGAGGG 116
DB 62 TACTCCGCTACAGACGCGCGGAGGAGTGAATGCTGGGGATCTGCGTAGAGAGGG 121
QY 117 GGTAAACGATCGAAGATGCTAATAACCGCATTAATCTTAAGAGGAAAGCAGGGGATC 176
DB 122 GATTAATCTAGTGAAGCGGTAGCTAATCCGATTAACGTCAGAACAGAGGGGAGCC 181
QY 177 GAAAGACCTTGGCGTTTGGAGCGCGGATGCTGATTAAGTGAATGTTGGGGTAAAGGC 236
DB 182 TTGGGGCTCTTGGCATCGATGAATGCCAGATGGGATTAAGCTAGTAGGGGTAAATGCG 241
QY 237 CTAAACAGGCGAGATGATGATGCTGAGAGAGACAGCAGCAGCACTGGAGCTAGA 296
DB 242 TCACTAGGCGAGCATCTTAAGCTGATGAGAGATGACCAAGCAGCACTGGAGCTAGA 301
QY 297 CAGGCGCAGACTCTTAAGGAGGAGCAGTGGGGAATTTTGAACAATGGCGCAGACCT 356
DB 302 CAGGCTCAGACTCTTAAGGAGGAGCAGTGGGGAATTTTGAACAATGGCGCAGACCT 361
QY 357 GATTCAGCAATGCCGCTGATGAAGAAGGCTTGGGTTGAAGCTCTTTCAGTGGAG 416
DB 362 GATGCAACCAATGCGCGTGTGTGAAGAAGGCTTGGGTTGAAGCACTTTCAGCAGAG 421
QY 417 AAGAAAGGTTAGCGTAAATATATGTCAGCCATGACCGATCGACAGAGAAAGACCGGC 476
DB 422 AAGAAAGCATTAAGTAAATTAATCTCGGTGATTAATCTCGAGAAAGAGACCGGC 481
QY 477 TAACTAGCGTCAGCAGCGCGGTAAATCTAGAGGTGACAGCGTAAATCGAATTAATCGG 536
DB 482 TAACTCGGTGACAGCGCGGTAAATCTAGAGGTGACAGCGTAAATCGAATTAATCGG 541
QY 537 GCGTAAAGGTGCGCAGCGCGCTTGAATGATGTAATCCCGGGCTTAACTCGG 596
DB 542 GCGTAAAGCGCAGCAGCGCGGTGTTGAATGTCGATGTAATCCCGGGCTCAACTCGG 601
QY 597 GAATTCGCTTGAAGCTCAAGAGCTAGAGTGGCAGAGGAGGTGAATTCATGTTGA 656
DB 602 GAATTCGCTTGAAGCTCAAGAGCTAGAGTGGCAGAGGAGGTGAATTCATGTTGA 661
QY 657 GCAGTGAATGCTAGAGATATGGAAGAATCATGATGCGAAGGAGCGCTCTGGGTAA 716
DB 662 GCGGTGAATGCTAGAGATATGGAAGAATCATGATGCGAAGGAGCGCGCTTGAAGA 721
QY 717 CACTGACGCTCATGACGAAAGCGTGGGAGCAAAAGAGATTAATCCCTGGTGTGCA 776
DB 722 GACTGACGCTCATGACGAAAGCGTGGGAGCAAAAGAGATTAATCCCTGGTGTGCA 781
QY 777 GCGCTTAACGATGTCATGATGTTGGGCTTATTAAGGCTTGG- TAAAGAAAGCTAAG 835
DB 782 GCGCTTAACGATGTCATGATGTTGGGCTTATTAAGGCTTGG- TAAAGAAAGCTAAG 841
QY 836 CGTGAAGTTGACCGCTGGGAGTACGCTGCGAAGATTAAACTCAAGAAATTGAACGAG 895
DB 842 CGTTAAGTTCGACCGCTGGGAGTACGCTGCGAAGATTAAACTCAAGAAATTGAACGAG 901
QY 896 GACCGGCAAGCGGTGATTAATGATGAATTAATTCGATGCAAGCGGAAAGAACTTAACTTA 955
DB 902 GCGCGGCAAGCGGTGAGTGTGTTAATGATGCAAGCGGAAAGAACTTAACTTA 961
```

```
QY 956 CCCTTGACATGTAAGCGAATTTTCTAGAGATTAAGTG- CTTGGGAAAGCTTAACAG 1014
DB 962 CTTTGACATGTAAGCGAATTTTCTAGAGATTAAGTG- CTTGGGAAAGCTTAACAG 1021
QY 1015 GTGCTGATGAGCTGCTGTCAGTCTGTCGATGATGTTGGTTAAGTCCCGCAAGAGC 1074
DB 1022 GTGCTGATGAGCTGCTGTCAGTCTGTCGATGATGTTGGTTAAGTCCCGCAAGAGC 1081
QY 1075 GCAACCTTGTCTAATTAATGTCATC- ATTTGGTGGGCACTTAAATGACATCCCGTGA 1133
DB 1082 GCAACCTTGTCTAATTAATGTCATC- ATTTGGTGGGCACTTAAATGACATCCCGTGA 1141
QY 1134 CAATCCGAGAGAGGTGGGAGTGAAGTCAAGTCTCAAGTCCCTTAATGAGTGGCTTCA 1193
DB 1142 TAAACTGAGAGAGGTGGGAGTGAAGTCAAGTCTCAAGTCCCTTAATGAGTGGCTTCA 1201
QY 1194 CAGTAAATCAATGAGCGCTGACAGAGGTTGCCAACCCCGAGGGGAGCTAATCTTCA 1253
DB 1202 CAGTAAATCAATGAGCGCTGACAGAGGTTGCCAACCCCGAGGGGAGCTAATCTTCA 1261
QY 1254 AAGCGGTGCTAGTCCGATCGAGTCTGCACTGCACTCCGTAAGTCCGAATCGTAG 1313
DB 1262 AAGTGGGTGCTAGTCCGATCGAGTCTGCACTGCACTCCGTAAGTCCGAATCGTAG 1321
QY 1314 TAATCCGATGACATGTCGCGGTGAATACCTTCCCGGTCTTGAACACAGCCCGTC 1373
DB 1322 TAATCCGATGACATGTCGCGGTGAATACCTTCCCGGTCTTGAACACAGCCCGTC 1381
QY 1374 AACCAATGAGAGGTGGGTTTCAACAGAGCAGTATGTTAACTGAAGAGGGCGCTTGC 1433
DB 1382 AACCAATGAGAGGTGGGTTTCAACAGAGCAGTATGTTAACTGAAGAGGGCGCTTGC 1441
QY 1434 ACGGTGATTCATGACTGGGGTG 1457
DB 1442 ACTTGTGATTCATGACTGGGGTG 1465

RESULT 19
US-10-831-286A-48666
; Sequence 48666, Application US/10831286A
; Publication No. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48666
; LENGTH: 1514
; TYPE: DNA
; ORGANISM: Cedecia davisiae
US-10-831-286A-48666

Query Match      69.0%; Score 1005.2; DB 11; Length 1514;
Best Local Similarity 82.2%; Pred. No. 5.1e-302;
Matches 1204; Conservative 0; Mismatches 253; Indels 7; Gaps 4;

QY 1 ATTGAACGCTGGGCGCATGCTTTACATGCAAGTGAAGGCGACGCGATGCTTGA- 59
DB 1 ATTGAACGCTGGGCGCGCCTTAACATGCAAGTGAAGGCGATGCAAGGATGCTTGC 60
QY 60 ---TCTGGTGGCGAGTGGCGGACGCGGTGAATGATCCGAAAGTATCCAGAGAGGG 116
DB 61 TACTCCGCTACAGACGCGCGGAGGAGTGAATGCTGGGGATCTGCGTAGAGAGGG 120
```



QY 357 GATCCGAATGCCGCTGAGTGAAGAGCCCTTCGGGTTGTAAGCTCTTTGAGTCGAG 416  
DB 361 GATGACAGCCATGCCGCTGATGAGAAAGAGCCCTTCGGGTTGTAAGCTCTTTGAGTCGAG 420  
QY 417 AAGAAAAGGTTACGGTAAATATATCGTGAACCATGACGGTATCGACAGAGAGAGACCGGC 476  
DB 421 AGAAGAGCGATGAGGTTAATTAATCACTGACAAATGATCTTACCGCAGAGAGAGACCGGC 480  
QY 477 TAACTACGTCGACGACCGCGGTAATACGTAGGGGTCAAGCGTTAATCGAAATTACG 536  
DB 481 TAACTCCGTCCGACGACCGCGGTAATACGTAGGGGTCAAGCGTTAATCGAAATTACG 540  
QY 537 GCGTAAAGGCTGCGCAGCGCGCTTGTAGTCAAGTGTGAATATCCCGGCTTAACCTGG 596  
DB 541 GCGTAAAGCGCAGCGCGGCTGTGTCAAGTGTGAATATCCCGGCTCAACCTGG 600  
QY 597 GAATTCGCTTTGAACCTACAAAGTATAGTGTGCAAGAGGAGGTGAATTCATGTGTGA 656  
DB 601 GAATTCGATTCGAAACTGGCAGGCTAGAGTCTGTGTAAGGGGGGTAGAATTCACGTTGTA 660  
QY 657 GCAGTGAATGCGTAGAGATATGGAAGAACATCGATGCGCAAGGACGCTCTGCGTTAA 716  
DB 661 GCGGTGAATGCGTAGAGATCTGGAAGAAATACCGGTGCGAAGGCGCGCTTGAGCAAA 720  
QY 717 CACTGACGCTCATGCAAGAAACGCTGGGAGCAAAAGATTAGATACCTGTGATGCCA 776  
DB 721 GACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAAGATTAGATACCTGTGATGCCA 780  
QY 777 GCGCTTAAAGATGTCAACTAGTGTGGCTTATTAAGCTTGG-71ACGAACTTAACG 835  
DB 781 GCGCTTAAAGATGTCAACTAGTGTGGCTTATTAAGCTTGGCTTCCGAACTTAACG 840  
QY 836 CGTGAAGTTGACCGCTGGGGAGTACGGTGCAGAGATTAAACTCAAGAAATTGACGGG 895  
DB 841 CGTGAAGTTGACCGCTGGGGAGTACGGTGCAGAGATTAAACTCAAGAAATTGACGGG 900  
QY 896 GACCCGCAAGACGGGTGATATGATTAATTGATGCAACGCGAAGAACTTACCTTA 955  
DB 901 GGCCCGCAAGACGGGTGAGCATGTGTTTAATTGATGCAACGCGAAGAACTTACCTTA 960  
QY 956 CCCTTGAATGATGCAAGTATTTCTAGAGATAGATTAGTG-CTTGGGAGACCTTAACAG 1014  
DB 961 CTCTTGAATGATGCAAGTATTTCTAGAGATAGATTAGTGCTTGGGAGACCTTAACAG 1020  
QY 1015 GTGCTGATGAGCTGCTGACGCTGTCGAGAGATGTTGGGTTAAGTCCGCAACGAGC 1074  
DB 1021 GTGCTGATGAGCTGCTGACGCTGTCGAGAGATGTTGGGTTAAGTCCGCAACGAGC 1080  
QY 1075 GCAACCCCTTGCATTAATTGCCATTC-ATTGTTGGGCACTTTAATGAGCTGCGGTGA 1133  
DB 1081 GCAACCCCTTATCTTTGTTGTCACGCGCTTAGCCGGGAACCTCAAGAGAGATGCGAAGTA 1140  
QY 1134 CAAGACCGAGAGAGGTGGGATGAGCTGCAAGTCTCATGSCCTTTATGGGTAGGGCTTCA 1193  
DB 1141 TAAATCGAGAGAGGTGGGATGAGCTGCAAGTCTCATGSCCTTTATCGAGTGGCTTCA 1200  
QY 1194 CAGCTAATATCAATGGCGGTACAGAGGTTTCCCAACCGCGAGGGGAGCTAATCTCGA 1253  
DB 1201 CACGTGCTACATGGCCGATACAAAGAACGACCTGCGAGAGCAAGCGAAGCTCTCTA 1260  
QY 1254 AAGGCGCTGATGTCGAGTCCGAGTCTGCAACTGACTCGTGAAGTCCGAATGCTGAG 1313  
DB 1261 AAGTGCCTGATGTCGAGTCCGAGTCTGCAACTGACTCGATGAAGTCCGAATGCTGAG 1320  
QY 1314 TAAATGCGAGATCAAGCATGTCCGAGTGAATACGTTCCGAGGTCTTGTACACACCGCCGTC 1373  
DB 1321 TAAATGCGAGATCAAGCATGTCCGAGTGAATACGTTCCGAGGTCTTGTACACACCGCCGTC 1380  
QY 1374 ACACCATGAGGTGGGTTTCAACAGAGAGGTATCTAACCGTGAAGAGGGCGCTTGGC 1433  
DB 1381 ACACCATGAGGTGGGTTTCAACAGAGAGGTATCTAACCGTGAAGAGGGCGCTTGGC 1440  
QY 1434 ACGGTGATTCATGACTGAGGCTG 1457

DB 1441 ACTTTGATTCATGACTGGGCTG 1464

RESULT 21  
US-10-831-286A-48678  
; Sequence 48678, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILDRES, DARRELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174, 001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; PRIOR FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 48678  
; LENGTH: 1535  
; ORGANISM: Proteus penneri  
US-10-831-286A-48678

Query Match 68.9%; Score 1003.8; DB 11; Length 1535;  
Best Local Similarity 82.6%; Pred. No. 1,4e-301;  
Matches 1210; Conservative 0; Mismatches 247; Indels 8; Gaps 5;

QY 1 ATGGAAGCTGGCGGAGCTGTTTACACATGCAAGTGCAGACGACAC--GATGCTTGC 58  
DB 21 ATTGAAGCTGGCGGAGCTGTTTACACATGCAAGTGCAGACGAGCGTAAAGAAAGCTTGC 80  
QY 59 AT--CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGGAACGTATCCAGAAAGAGGG 116  
DB 81 TTTCTTGCTTACAGACGCGCGGAGCGGTGAGTAATGTAATGGGATCTGCCGATAGAGGG 140  
QY 117 GGTAGCGCATGGAAGATGTGCTAATATACCGCATATACTCTAAGAGAGGAAAGCAGGGGATC 176  
DB 141 GATTACTACTGGAAGACGCTGATTAATACGCAATGACGCTTACGGAACCAAGAGGGGCTC 200  
QY 177 GAAAGACCTTGCCTTTTGGAGCGCGCATGTCTGATTAAGTCTAGTGGGTAAAGGC 236  
DB 201 TTGGGACCTTGGCGTATCGGATGAACCATATGAGATTAAGTCTAGTGGGTAAAGGC 260  
QY 237 CTACCAAGGCGAGTATGATGTTGCTGAGAGAGACACAGCCACACTGGGACTGAGA 296  
DB 261 TCACCTAGGCGAGATCTTATGCTGTGAGAGAGATCAAGCACACTGGGACTGAGA 320  
QY 297 CACGCGCCACATCTCTACGAGGAGGACAGCATGGGGAAATTTTGGCAATGGGCGCAAGCT 356  
DB 321 CACGCGCCACATCTCTACGAGGAGGACAGCATGGGGAAATTTTGGCAATGGGCGCAAGCT 380  
QY 357 GATCCAGCAATGCGCGCTGAGTGAAGAGGCTTCCGGTTGTAAAGCTCTTTCACTGAG 416  
DB 381 GATCCAGCAATGCGCGCTGATGAAGAGGCTTAAAGAGGCTTAAAGTATCTTCAAGCGGG 440  
QY 417 AAGAAAAGTTACCGTAAATATATCGTGAACCATGACGGTATCGACAGAGAGACCGGC 476  
DB 441 AGAAGGTGATTAAGTTAATATCTTTATCAATTGACCTTACCCGCAAGAGAGACCGGC 500  
QY 477 TAACTACGTCGACGACCGCGGTAATACGTAGGGGTCAAGCGTTAATCGGAATTACG 536  
DB 501 TAACTCCGTCCGACGACCGCGGTAATACGTAGGGGTCAAGCGTTAATCGGAATTACG 560  
QY 537 GCGTAAAGGCTGCGCAGCGCGCTTGTAGTCAAGTGTGAATATCCCGGCTTAACCTGG 596  
DB 561 GCGTAAAGCGCAGCGAGCGGTGCAATTATGATGATGTAAGAGCCCGGAGCTTAACCTGG 620  
QY 597 GAATTCGCTTTGAACCTACAAAGTATAGTGTGCAAGAGGAGGTGAATTCATGTGTA 656



DB 621 GAATTCATCTGAACCTGGTGGCTAGAGCTTGTGAGAGGGGGGTGAATTCATCTGTA 680  
QY 657 GCAGTGAATGCGTAGAGATATGAAAGAACATGATGCGAGCGACCTCTGGGTAA 716  
DB 681 GCGGTGAATGCGTAGAGATATGTAAGAAATACCGGTGGCGAAGCGCCCTGGAGCAAA 740  
QY 717 CACTGACGCTCAATGACGAAAGCGTGGGAGCAAAAGGATTGATACCTGGTAGTCCA 776  
DB 741 GACTGACGCTCAGGTGGGAGCGTGGGAGCAAAAGGATTGATACCTGGTAGTCCA 800  
QY 777 CGCCCTAAACGATGTC-AACTAGTGTGGGCTTATTAAGCTTGGTAAAGAACTAAG 835  
DB 801 CGCTGTAACGATGTGATTTAGAGTGTGTGCTTGAACCGGTGCTTGGAGCTAAG 860  
QY 836 CGTGAAGTTGACCGCTGGGAGTACCGGTGCAAGATTAAATCTAAAGAAATTGACGG 895  
DB 861 CGTTAATGACCGCTGGGAGTACCGCGCAAGGTTAAACTAAATGAATTGACGGG 920  
QY 896 GACCCGCAACAGCGGTGATTAATGATTAATTCATGCAAGCGCAAAACCTTACCTA 955  
DB 921 GGGCCGCAACAGCGGTGATTAATGATTAATTCATGCAAGCGCAAAACCTTACCTA 980  
QY 956 CCCTTGACATGATCGCAATTTTCTAGAGATAGATTAGTG-CTTGGGAAAGCTAACAAG 1014  
DB 981 CTTTGAATCAGCGGATCTTTAGAGATAGAGAGTGGCTTCGGGAACGTGAGACAG 1040  
QY 1015 GTGCTGATGCTGTGCTGACGCTGTGTCTGATGATGTTGGTTAAAGTCCGCAACGAGC 1074  
DB 1041 GTGCTGATGCTGTGCTGACGCTGTGTCTGATGATGTTGGTTAAAGTCCGCAACGAGC 1100  
QY 1075 GCAACCTTGTCAATTAATGATGATGATG-GTGGGCACTTAAAGACCTGCGGCTG 1132  
DB 1101 GCAACCTTGTCAATTAATGATGATGATG-GTGGGCACTTAAAGACCTGCGGCTG 1160  
QY 1133 ACAACCGGAGAAAGTGGGAGTACGTCAGATCCTCATGAGCCCTTATGGGTAGGGCTTC 1192  
DB 1161 ATTAACCGGAGAAAGTGGGAGTACGTCAGATCCTCATGAGCCCTTATGGGTAGGGCTTC 1220  
QY 1193 ACAAGTATCAATGAGCGCTACAGAGAGGTTGCCAACCGCGAGGGGAGCTAATCTCAG 1252  
DB 1221 ACAGGTGCTCAATGAGCGCTACAGAGAGGTTGCCAACCGCGAGGGGAGCTAATCTCAG 1280  
QY 1253 AAAGCGGTGATGCTGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1312  
DB 1281 AAAGTGTGTGATGCTGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1340  
QY 1313 GTAATCGCGATGAGCATGTCGCGGTGATACGTTCCGGGTCTTGTACACACCGCCGT 1372  
DB 1341 GTATCGTATGATGATGATGCTACGCTGATACGTTCCGGGTCTTGTACACACCGCCGT 1400  
QY 1373 CACACCATGAGAGTGGGTTTACACAGAAAGAGTATGTAACCGTAAAGAGGCGCTTGC 1432  
DB 1401 CACACCATGAGAGTGGGTTTACACAGAAAGAGTATGTAACCGTAAAGAGGCGCTTGC 1460  
QY 1433 CACGCTGAGATTCTATGATCTGGGTG 1457  
DB 1461 CACTTGTGATTCATGACTGGGGTG 1485

RESULT 22  
US-10-831-286A-48670/c  
; Sequence 48670, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILDRESS, DARRYL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A

; CURRENT FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; PRIOR FILING DATE: 2003-04-24  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 48670  
; LENGTH: 1501  
; TYPE: DNA  
; ORGANISM: *Cedecea neteri*  
US-10-831-286A-48670  
Query Match 68.9%; Score 1003.6; DB 11; Length 1501;  
Best Local Similarity 82.2%; Pred. No. 1.6e-301;  
Matches 1203; Conservative 0; Mismatches 254; Indels 7; Gaps 4;  
QY 1 ATTTGAACGCTGGCGGCATGCTTTTACATGCAATGCAAGTCGAAACGGCAGACGAGATGCTTG--- 57  
DB 1497 ATTTGAACGCTGGCGGCAGGCTTAAACATGCAAGTCGAAACGGGTAGACCGGAGAGCTTGCC 1438  
QY 58 -CATCTGGTGGAGTGGCGGAGCGGAGTGAATGATCATCGAAACGTATCCAGAAAGGGG 116  
DB 1437 TCTCTGGTGAAGACGCGGAGCGGAGCGGAGTGAATGATCTGGGATGCTGCTGATGAGGGG 1378  
QY 117 GGTAAACGATCGAAAGATGCTAAATACCGCATATCTCTAAAGAGAAAGCAAGGGGATC 176  
DB 1377 GATTAATGAGAAACGATGATGCTAAATACCGCATATCTCTAAAGAGAAAGCAAGGGGATC 1318  
QY 177 GAAAGACCTTGGGCTTTTGGAGCGGCGCATGCTGATTAAGTATGTTGGGTAAAGGC 236  
DB 1317 TTGCGGCTCTTGGCATGATGAAACCGAGATGGAATTAAGTGAAGTGAATGATGCT 1258  
QY 237 CTACCAAGGCGAGATGATGATGTTGGTCTGAGAGACGACCACTGAGGACTGAGA 296  
DB 1257 TCACCTAGGCGAGATGCTTACCTGATGCTGATGAGATGACGACCACTGAGACTGAGA 1198  
QY 297 CACGCGCCAGACTCTTACGCGAGGAGCAGAGCTGGGAAATTTGAGCAATGCGCGCAAGCT 356  
DB 1197 CACGCGCCAGACTCTTACGCGAGGAGCAGAGCTGGGAAATTTGAGCAATGCGCGCAAGCT 1138  
QY 357 GATCCAGCAATGCGCGGTGATGAAAGAGCCTTGGGTTGAAGGCTTTCAATGCGAG 416  
DB 1137 GATGCGCATGCTCGGCTGATGAAAGAGCCTTGGGTTGAAGGCTTTCAATGCGAG 1078  
QY 417 AAGAAAGGTTACGGTAAATTAATCGTACCGTACCGTATGAGAGGCTTAATGGAATTA 476  
DB 1077 AAGAAAGGTTACGGTAAATTAATCGTACCGTATGAGAGGCTTAATGGAATTA 1018  
QY 477 TAACTACGTGACGAGCGCGGTAAATGTAAGGCTGCAAGGCTTAATGGAATTA 536  
DB 1017 TAACTACGTGACGAGCGCGGTAAATGTAAGGCTGCAAGGCTTAATGGAATTA 958  
QY 537 GCGTAAAGGCTGCGAGCGGCGCTTGTAGTCAAGTGAATTCGCCGGGCTTAACCTG 596  
DB 957 GCGTAAAGGCTGCGAGCGGCGCTTGTAGTCAAGTGAATTCGCCGGGCTTAACCTG 898  
QY 597 GAATTCGTTTGAATTAATCAAAAGCTAGAGTGTGCGAGAGGAGTGAATTCATGCTGA 656  
DB 897 GAATTCGTTTGAATTAATCAAAAGCTAGAGTGTGCGAGAGGAGTGAATTCATGCTGA 838  
QY 657 GCAGTGAATGCGTAGAGATATGAAAGACATGATGCGAAAGCGCTCTGGGTAA 716  
DB 837 GCGGTGAATGCGTAGAGATATGAAAGACATGATGCGAAAGCGCTCTGGGTAA 778  
QY 717 CACTGACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAAGTAACTCGTGTAGTCA 776  
DB 777 GACTGACGCTCATGAGTGCAGAAAGCGTGGGAGCAAAACAGATTAAGTAACTCGTGTAGTCA 718  
QY 777 CGCCCTAAACGATGTAATGATGTTGGGCTTATTAAGCTTGG-TAAAGAACTAAG 835  
DB 717 CGCCCTAAACGATGTAATGATGTTGGGCTTATTAAGCTTGG-TAAAGAACTAAG 658  
QY 836 CGTGAAGTTGACCGCTGGGAGTACGCTGCAAGATTAAACTCAAGAAATTGACGGG 895



Db 657 CGTTAAGTCGACCGCCTGGGGAGTAGCGCGCAAGTTAAACTCAAAATGATGACGG 598  
Qy 896 GACCCCGCAAGCGGTGGATTATGTGATTAATTCATGCAACCGAAAACTTACTTA 955  
Db 597 GGCCCGCAAGCGGTGGAGCATGTGTTAATTCATGCAACCGAAGAACTTACTTA 538  
Qy 956 CCCTTGACATGTAGCCAAATTTTCTAGATATGATTACTG-CCTCGGGAAGCTTAACAG 1014  
Db 537 CTCTTGACATGTAGCAAGAACTTTCCAGATGTAGTTGTGCTTGGGAACCTGAGACAG 478  
Qy 1015 GTGCTGATGGCTGTGCTGACGCTGTGTCGATGATGTTGGGTTAAGTCCCGCAACGAGC 1074  
Db 477 GTGCTGATGGCTGTGCTGACGCTGTGTTGGAAATGTTGGTTAAGTCCCGCAACGAGC 418  
Qy 1075 GCAACCCCTTGATTAATGTCATC-ATTGTTGGGCACTTAAATGATGATGCTCCGCTGA 1133  
Db 417 GCACCCCTTAATCTTTGTGTCAGCGGTTGGCGGAACTCAAGAGACCTGCACTGA 358  
Qy 1134 CAACCCGAGAGAGTGGGATGACGTCAGTCTCTCATGCGCCCTTAATGAGTGGGCTTCA 1193  
Db 357 TAAACTGAGAGAGAGTGGGATGACGTCAGTCTCATGAGCCCTTAAGAGTGGGCTTCA 298  
Qy 1194 CACCTAATACAGTGGCGGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTGAGA 1253  
Db 297 CACGTGCTACATGCGCATACAAAGAGACGACCTGCGAGAGCAAGCGGAACTCATTA 238  
Qy 1254 AAGCGCTGTAGTCCGGATCGAGTCTGCAACTGCACTCCGTAAGTCCGAATCGCTAG 1313  
Db 237 AAGTGGCTGTAGTCCGGATCGAGTCTGCAACTGCACTCATGAGTCCGAATCGCTAG 178  
Qy 1314 TAATCGCGAGTACAGATGTCGCGGTGAATACGTTCCCGGCTTGTATACACCGCCGCTC 1373  
Db 177 TAATCGTAGATCAAGAAATGCTACGGTGAATACGTTCCCGGCTTGTATACACCGCCGCTC 118  
Qy 1374 ACACCATGAGAGTGGTTCACCAAGACAGTAGTCTAACCGTAAGAGGAGCGCTTACC 1433  
Db 117 ACACCATGAGAGTGGTTCACAAAGAGAGTAGTCTAACCTTCGGAGGAGCGCTTACC 58  
Qy 1434 ACGGTGATTCATGACTGGGATG 1457  
Db 57 ACTTGTGATTCATGACTGGGATG 34

RESULT 23  
US-10-831-286A-48693  
; Sequence 48693, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILDRESS, DARRELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; PRIOR FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; PRIOR FILING DATE: 2003-04-24  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 48693  
; LENGTH: 1527  
; TYPE: DNA  
; ORGANISM: Morganella morganii sp. morganii  
US-10-831-286A-48693

Query Match 68.8%; Score 1002.8; DB 11; Length 1527;  
Beef Local Similarity 82.5%; Pred No. 2.8e-301;  
Matches 1210; Conservative 0; Mismatches 247; Indels 9; Gaps 5;

Qy 1 ATTGAACGCTGCGGACATGCTTTA-CACATGCAAGTCCGAACGAGCAGACGGA---TGCT 55  
|||||

Db 14 ATTGAACGCTGCGGACGACCTTAAGCATGCAAGTCCGCGGTAAACAGGAGAACCTTG 73  
Qy 56 TGCACTGGTGGGAGAGTGGCGGACGSGTGAATGATCATCGAACCTATACAGAAAGAG 115  
Db 74 CTCTCTGCTGACAGACCGGCGAGCGGTGATATATAGGGATCTGCTGATAGCGGG 133  
Qy 116 GGGTAACGATCCGAAGATGTGTAAATACCGCATATATCTTAAGAGGAAAGCAGGGAT 175  
Db 134 GGAATTAATCTGGAAGACGATTAATACCGCATATATGCTTCGACCAAGACGCGGGAGC 193  
Qy 176 CGAAGACCTTGGCTTTTGGAGCGGCCGATGCTCTGATTAAGCTATGTTGGTGGGTTAAAG 235  
Db 194 CTCGGGCGCTCGCGCATCATGATGAACCATATAGGATTAAGCTATGAGTGAAGTTAACGG 253  
Qy 236 CTTACCAAGCGCGACATGATGATGTTGGTGAAGAGACGACGACACTGGAGACTGAG 295  
Db 254 CTTACCTAAGCGACATGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 313  
Qy 296 ACAAGCGCGGACGCTCTTACGAGAGGACGAGTGGGGAATTTTGAACATGAGGCGCAAGCC 355  
Db 314 ACAAGCGCGGACGCTCTTACGAGAGGACGAGTGGGGAATTTTGAACATGAGGCGCAAGCC 373  
Qy 356 TGATCCAGCAATCCGCGGTGATGAAGAGGCTTCGCGTTGTAAGCTCTTCACTGCA 415  
Db 374 TGATGAGCGCATCGCGGTGATGAAGAGGCTTCGCGTTGTAAGCTCTTCACTGCG 433  
Qy 416 GAAAGAAAGGTTACGTTAATATCGTGAACCCATGACGGTATGACAGAGAGAGACACCG 475  
Db 434 GAGGAGAGTGGCAAGGTTAATACCTTGGCAATGACGTTACGACAGAGAGAGAGACACCG 493  
Qy 476 CTAATCAAGTCCAGACGCGCGGTATATAGTAAAGGTCGAAGGTTAATCGAATTTACTG 535  
Db 494 CTAATCTCGTCCAGACGCGCGGTATATAGGAGGTCGAAGGTTAATCGAATTTACTG 553  
Qy 536 GCGTAAAGGAGTGGCGAGCGCGCTTGAATGATGATGAATTCCTCGGCTTAACTG 595  
Db 554 GCGTAAAGGAGCGAGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 613  
Qy 596 GGAATTCGTTGAACTCAAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 655  
Db 614 GGAATTCATCTGATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 673  
Qy 656 AGCATGAATTCGATGAATGAGAGAAATCATGATGAGGAGGAGGAGGAGGAGGAGGAGG 715  
Db 674 AGCGTGAATTCGATGAATGAGAGAAATCATGATGAGGAGGAGGAGGAGGAGGAGGAGG 733  
Qy 716 ACACTGACGCTCATGACGAAGAGGTCGGAGGAGCAACAGATTAAGATACCTGCTAGTCC 775  
Db 734 AGACTGACGCTCAGAGTGCAGAAAGCTGGGAGAGCAACAGATTAAGATACCTGCTAGTCC 793  
Qy 776 ACGGCTTAACGATGCACTAGTGTGAGGCTTATGAGCTTGG-TAACGAAGCTAAC 834  
Db 794 ACGGCTTAACGATGCACTAGTGTGAGGCTTATGAGGCTTGGAGGCTTGGAGGCTTGGAGG 853  
Qy 835 GCGTGAAGTGAACCGCTCGGGAGTACGTCGCAAGATTAACATCAAGGAATTTGACGG 894  
Db 854 GCGTGAAGTGAACCGCTCGGGAGTACGTCGCAAGATTTAAACTCAAGTAATTTGACGG 913  
Qy 895 GGAACCGGACAAAGCGGTGATTAATGATTAATTTGATGATGATGATGATGATGATGATGATG 954  
Db 914 GGGCCCGACAAAGCGGTGAGAGATGTGTTAATTCATGACGAGCAAGAACTTTACT 973  
Qy 955 ACCCTTGACATGATGCAATTTTCTAGAGATGATGATGATGATGATGATGATGATGATGATG 1013  
Db 974 ACTCTTGACATGCAAGAGACTTGGCAAGATGCTTGTGTGCTTGGGAACTTGGAGCA 1033  
Qy 1014 GTGCTGATGAGTGTGTGCTGACGCTGTGTGTGATGATGATGATGATGATGATGATGATGATG 1073  
Db 1034 GTGCTGATGAGTGTGTGCTGACGCTGTGTGTGATGATGATGATGATGATGATGATGATGATG 1093  
Qy 1074 CGCAACCTTGTCTAATTAATGCAATTTG--GTTGGGCACTTAATGAGTCCGGT 1131  
Db 1094 CGCAACCTTATCTTGTGTGCAAGCGGTGATGAGGAGAACTCAAGAGAGACTGCGGT 1153

QY 1132 GACAAACCGGAGAGAGTGGGGATGACGTCAAGTCTCTCATGCGCTTATGGTAGGGCTT 1191  
| | | | |  
DB 1154 GATTAACCGGAGAGAGTGGGGATGACGTCAAGTCTCTCATGCGCTTATGGTAGGGCTT 1213  
| | | | |  
QY 1192 CACAGCTTAATACAAATGAGCGCGTACAGAGGGTTGCCAACCCCGGAGGGGGAGCTTAATCTCA 1251  
| | | | |  
DB 1214 CACAGCTCTACAAATGAGCGCGTACAGAGGGATCGAACCCCGGAGGGGGAGCTTAATCTCA 1273  
| | | | |  
QY 1252 GAAAGCGCGTCTGATGTCGGATCGAGTCTGCAACTCGACTCCGTGAAGTGGGAATCGCT 1311  
| | | | |  
DB 1274 TAAAGTACGTCTGATGTCGGATCGAGTCTGCAACTCGACTCCGTGAAGTGGGAATCGCT 1333  
| | | | |  
QY 1312 AGTAATCGCGAGTACAGATCTCCCGGTGAATACGTTCCCGGTCTTGTACACACCCGCCG 1371  
| | | | |  
DB 1334 AGTAATCGTGAATCAGAAATCTACCGGTGAATACGTTCCCGGTCTTGTACACACCCGCCG 1393  
| | | | |  
QY 1372 TCACACCATGGAGTGGGGTTTCCACAGAACGAGTACTTAACCGTAAGAGGGGGCGCTTG 1431  
| | | | |  
DB 1394 TCACACCATGGAGTGGGGTTTCCACAGAACGAGTACTTAACCGTAAGAGGGGGCGCTTG 1453  
| | | | |  
QY 1432 CCACGCTGAGATTCACTACTGGGGTG 1457  
| | | | |  
DB 1454 CCACTTTGTGATTCATGACTGGGGTG 1479  
| | | | |  
RESULT 24  
US-10-831-286A-48679  
/ Sequence 48679, Application US/10831286A  
/ Publication No. US20060046246A1  
/ GENERAL INFORMATION:  
/ APPLICANT: ZENG, QIANDONG  
/ APPLICANT: CHATELLIER, SONIA  
/ APPLICANT: MOIR, DONALD T.  
/ APPLICANT: LACROIX, BRUNA  
/ APPLICANT: CHILDRESS, DARRELL  
/ TITLE OF INVENTION: GENUS, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
/ FILE REFERENCE: 032796-174.001  
/ CURRENT APPLICATION NUMBER: US/10/831.286A  
/ PRIOR FILING DATE: 2004-04-26  
/ PRIOR APPLICATION NUMBER: 60/464,955  
/ NUMBER OF SEQ ID NOS: 48788  
/ SOFTWARE: Patentin version 3.2  
/ SEQ ID NO 48679  
/ LENGTH: 1486  
/ TYPE: DNA  
/ ORGANISM: Enterobacter hormaechei  
US-10-831-286A-48679  
Query Match 68.6%; Score 1000; DB 11; Length 1486;  
Best Local Similarity 82.3%; Pred. No. 2.1e-300;  
Matches 1197; Conservative 0; Mismatches 250; Indels 7; Gaps 4;  
QY 11 GGGCGCATGCTTTACATGCAAGTCCGACGCGACGCGA----TGCCTTGCAATCTGTG 66  
| | | | |  
DB 1 GGGCGCGAGGCTTAACATGCAAGTCCGACGCGTAACGGAAGCAGCTTGCTGCTTGTGCTG 60  
| | | | |  
QY 67 GCGAGTGGCGGACGGGTGAGTATGATCGGAACGTATCCAGAAAGAGGGGGTAAACGCAT 126  
| | | | |  
DB 61 ACGAGTGGCGGACGGGTGAGTATGATCTGCGAAACCTGCTGATGGAGGGGATTAACCTACT 120  
| | | | |  
QY 127 CGAAAGATGTGCTAATACCGCATATACTCTTAAGAGGAAAGACAGGGGATCGAAAGACTT 186  
| | | | |  
DB 121 GGAACCGTGTATATACCGCATATACGTCCGAAAGACCAAGAGGGGGACCTTCGCGGCTC 180  
| | | | |  
QY 187 GCGCTTTTGAAGCGCGCGATCTCTGATTAAGTATGTTGGTGGGTAAGGGCTTACCAAGGC 246  
| | | | |  
DB 181 TTGCAATCGGATGTGCCAAGATGGGATTAAGTAAAGTGGGTAACGCTCACTTAGGC 240  
| | | | |  
QY 247 GAGCATAGTATGTTCTGAGAGGACGACGACCACTGGGACTGAGACACGCGCCAG 306  
| | | | |  
DB 241 GAGCATCCCTTAGTGTCTGAGAGGATGACCAAGCCACACTGGAACTGAGACACGCTCAG 300  
| | | | |

QY 307 ACTCTACGGGAGGACGAGCAGTGGGAATTTTGAACAATGGCGCAAGCCTGATCCAGCAA 366  
| | | | |  
DB 301 ACTCTACGGGAGGACGAGCAGTGGGAATTTTGAACAATGGCGCAAGCCTGATCCAGCAA 360  
| | | | |  
QY 367 TCCCGGTGAGTGAAGAAAGCCTTCGGGTGTAAGAGCTCTTTCACTCCAGAAAGAAAGT 426  
| | | | |  
DB 361 TCCCGGTGATGAAGAAAGCCTTCGGGTGTAAGAGCTCTTTCACTCCAGAAAGAAAGT 420  
| | | | |  
QY 427 TACGTAATTAATCCGTAACCCATGACGATGACAGAAAGAAAGCAACCGGCTAACTAGCTG 486  
| | | | |  
DB 421 TAAGTTAATTAATCTTGTGATTAAGCTTAACCGAGAAAGAAAGCAACCGGCTAACTAGCTG 480  
| | | | |  
QY 487 CCAGACCGCGCGTAAATAGTAGGGTGCAGCGTTAATCGAATTAATCTGGCGTAAAGGG 546  
| | | | |  
DB 481 CCAAGACCGCGCGTAAATAGTAGGGTGCAGCGTTAATCGAATTAATCTGGCGTAAAGGG 540  
| | | | |  
QY 547 TCCGACGCGCGCTTGTAAAGTGAATGGAATCCCGGGCTTAACTTGGGAATTTGCGTT 606  
| | | | |  
DB 541 CACGACGCGCGCTTGTCAAGTGGATGTGAATCCCGGGCTCAACTTGGGAATTTGCAAT 600  
| | | | |  
QY 607 TGAATCTACAAAGCTTAGAGTGTGCGAGAGGAGTGAATTCATGTTAGAGTGAAT 666  
| | | | |  
DB 601 CCAAACTGGCAGCTTAGAGTCTTGTAGAGGGGGTGAATTTCAAGTGTAGCGGTGAAT 660  
| | | | |  
QY 667 GCGTAGAGTATGGAAGAACATCGATGGCGAAGCAGCCTCTGGGTTAACACTGACGCT 726  
| | | | |  
DB 661 GCGTAGAGTATGGAAGAACATCCGTTGGCGAAGCGCGCCCTTGAACAAGACTGACGCT 720  
| | | | |  
QY 727 CATGACGAAACGTTGGGGAGGAAACAGGATTAATCCTGTGTGTCCAGCCCTTAAC 786  
| | | | |  
DB 721 CAGGTCCGAAACGTTGGGGAGGAAACAGGATTAATCCTGTGTGTCCAGCCCTTAAC 780  
| | | | |  
QY 787 GATGTCACTAGTGTGGGCTTATTAAGCTTGG- TAAAGAGCTAACGCGTGAAGTTG 845  
| | | | |  
DB 781 GATGTCACTAGTGTGGGCTTATTAAGCTTGG- TAAAGAGCTAACGCGTGAAGTTG 840  
| | | | |  
QY 846 ACCGCTGGGAGTAGTGGTTCGCAAGATTAACTCAAAAGAAATGACGGGGACCCGAC 905  
| | | | |  
DB 841 ACCGCTGGGAGTAGTGGTTCGCAAGATTAACTCAAAAGAAATGACGGGGACCCGAC 900  
| | | | |  
QY 906 AGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAACCTTAACCTTGTGACAT 965  
| | | | |  
DB 901 AGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAACCTTAACCTTGTGACAT 960  
| | | | |  
QY 966 GTAGCAATTTTCTAAGATTAAGTATAGT-CTTCGGGAACGCTAACACAGGTGTCGATG 1024  
| | | | |  
DB 961 CCAAGAACTTAAGCAAGATGCTTGTGCTTCGGGAACCTTAAGACAGGTGTCGATG 1020  
| | | | |  
QY 1025 GCTGTGTCAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAACGACGCAACCCCTTG 1084  
| | | | |  
DB 1021 GCTGTGTCAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAACGACGCAACCCCTTA 1080  
| | | | |  
QY 1085 TCATTAATTTGCATC-ATTTGTTGGGCACTTAAATGAGACTGCGGTGACAAACCGGAG 1143  
| | | | |  
DB 1081 TCCCTTTGTTGCAAGCGGTTAAGCGCGGAACCTCAAGAGAGACTGCGATTAACCTGAGG 1140  
| | | | |  
QY 1144 GAAAGTGGGATGACGTCAAGTCTCATGCGCTTATGGGTAGGGCTTCAACGTAATAC 1203  
| | | | |  
DB 1141 GAAAGTGGGATGACGTCAAGTCTCATGCGCTTATGGGTAGGGCTTCAACGTCATAC 1200  
| | | | |  
QY 1204 AATGGCGGCTAACAAGAGGTTGCCAACCGCGAGGGGAGCTAATCTCAGAAAGCGGCTG 1263  
| | | | |  
DB 1201 AATGGCGGCTAACAAGAGAGGACCTCCGAGAGACAGCGACCTCAATTAAGTGTGCTG 1260  
| | | | |  
QY 1264 TAGTCCGATCGAGTCTGCAACTCGACTCCGTGAATCGGAATCGTAATTCGCGGA 1323  
| | | | |  
DB 1261 TAGTCCGATCGAGTCTGCAACTCGACTCAATGAATCGGAATCGTAATTCGCGGA 1320  
| | | | |  
QY 1324 TCAGCATGTGCGGTGAATACGTTCCGGGTCTTGTACACACCGCCGCTACACCATGGG 1383  
| | | | |  
DB 1321 TCAGAAATGCCACGTTGAATACGTTCCGGGCTTGTACACACCGCCGCTACACCATGGG 1380  
| | | | |



```

Db      1288 AAGTGGCGTCGTAAGCCGGATTGAGATCGCAACTGCACTCCATGAAAGTCGGAATCGGTAG 1347
Qy      1314 TAAATCGCGGATCAAGCATGTCCGCGGTGAATACGTTCCGGGTTCTTTGACACACCGCCGCTC 1373
Db      1348 TAAATCGTGAATCAGAAATGCCACCGGTGAATACGTTCCCGGGCTTTGTACACACCGCCGCTC 1407
Qy      1374 ACACCATGGGAGTGGGTTTCCACGAACAGAGTGTCTACACCGTAAGAGGAGCGCTTGCC 1433
Db      1408 ACACCATGGGAGTGGGTTTCCAAAGAAAGTAAGTACTTAACTTTCGGAGGAGCGCTTAAC 1467
Qy      1434 ACGGTGATTCATGACTCGGAGTG 1457
Db      1468 ACTTTGTCATTCATGACTCGGAGTG 1491

RESULT 26
US-11-198-746-158
Sequence 158, Application US/11198746
Publication No. US20060035256A1
GENERAL INFORMATION:
APPLICANT: BROW, MARY ANN D.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: OLIVE, DAVID M.
TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
TITLE OF INVENTION: PATHOGENS
NUMBER OF SEQUENCES: 165
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/198, 746
FILING DATE: 05-Aug-2005
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/941,193
FILING DATE: 28-Aug-2001
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FOS-01756
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 158:
SEQUENCE CHARACTERISTICS:
LENGTH: 1542 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-11-198-746-158

Query Match      68.4%; Score 996.8; DB 18; Length 1542;
Best Local Similarity 82.3%; Pred. No. 2,1e-299;
Matches 1205; Conservative 0; Mismatches 252; Indels 7; Gaps 5;

Qy      59 ATC--TGCTGGGAGATGGCGGACGGGTGAGTAATGATCGGAACGATACAGAAAGAGGG 116
Db      28 ATTTGAACGCTGGCGGACGAGCTTAACACATGCAAGTCGAACGGTTAACAGGAACACTTGC 87
Qy      88 TTCCTTTCGACGAGTGGCGGACGGGTGATGATGCTGGGAAACTCCTGATGAGAGGG 147

```

QY	117	GGTACGCATCGAAGAATGTGCTAAATCCGCAATACTCTAAGAGAGAAAGCGGGATC	176
Db	148	GATTAACATACGAGAAACGGTAGCTAAATACCGCATATACTCCCAAGACCAACAGAGGGAGACC	207
QY	177	GAAGAAGCCTTGCGCTTTTGGAGCGGCGGAGTGTCTGATTAGCTAGTTGGTGGGTAAAGGC	236
Db	208	TTGCGGCGCTCTTGCGATCGGATGTGCCAAGATGGGATTTACTAGTAGTGGGGTAAAGGC	267
QY	237	CTACCAAGCGGACGATCAGTAGTTGGTCTGAGAGAGACGACACAGCACACTGGGACTGAGA	296
Db	268	TCACCTTAGCGGACGATCCCTAGCTGTGCTGAGAGAGTAGACACGACACACTGGAACCTAGA	327
QY	297	CACGGCCCAACTCTTAACGGGAGGACACACTGGGGAAATTTTGGACATGGGCGCAAGCT	356
Db	328	CACGGTCCAGCTCTTAACGGGAGGACAGTGGGGAAATTTGCAAAATGGGCGCAAGCT	387
QY	357	GATCCAGCAATGCCGCGTAGTGAAAGAAAGGCTTCCGGTTGTAAAGCTCTTTCAGTGGAG	416
Db	388	GATCCAGCCATGCCGCGTAGTAGAAGAAAGGCTTCCGGTTGTAAAGTACTTTCAGCGGGG	447
QY	417	AAGAAAAGTTACGGTAAATTAATGTGAGACCATGACGGTATTCGACAGAGAGACCGGC	476
Db	448	AGGAAGGAGTAAAGTTAATTAATCTTTGCTCATTAACGTTACCCGCAAGAGACCGGC	507
QY	477	TAATCTAGCTGCACAGACCGCGGTAAATACGTAGGGTCCAGACGCTTAATCCGAATTAATCG	536
Db	508	TAACTCCGTGCACACACCGCGGTAAATACGAGGGGTCCAGCGCTTAATCCGAATTAATCG	567
QY	537	GGGTAAAGGGGCGCGCAGGCGGCTTGTAAAGTCAGATGTGAATCCCGGGCTTAATCCCTGG	596
Db	568	GGGTAAAGCGCACGACGCGGTTTGTAAAGTCAGATGTGAATCCCGGGCTCAACTGG	627
QY	597	GAATTTGGCTTGAACCTACAAAGCTAGAGTGTGGCAGAGGGAGGTGGAATTCATGTGTA	656
Db	628	GAACTGCATCTGAAATCTGGCAAGCTTAGTCTGTAGAGGGGGGTGAATTCACAGTGTA	687
QY	657	GCAGTGAATTCGTAGAGATATGGAAGAACATGCAATGGCCGAAGGCAACCTCTGGGTAA	716
Db	688	GGGTGAATTCGTAGAGATCTGAGAGAAATACCGTGTGGCAAGGCGGCCCTGTGACGAA	747
QY	717	CACAGCGCTCATGCAAGAAAGCGTGGGGAGCAAAAGAAATTAATACCTCGTAGTCCA	776
Db	748	GACTGACGCTCAGGTGCGAAAGCGTGGGGAGCAAAAGAAATTAATACCTCGTAGTCCA	807
QY	777	CGCCCTAAACGATGTCAACTAGTTGTGGGCTTATTAAGCTTTGG-TAAAGAACTAACG	835
Db	808	CGCCGTAAACGATGTCAACTAGTTGTGGGCTTATTAAGCTTTGGGCTTCCGAACTAACG	867
QY	836	CGTAGAAGTTGACGCGCTGGGGAGTACCGTGCAGCAGATTAATACTCAAGAAATTTGACGGG	895
Db	868	CGTTAAGTCGACCGCTGGGGAGTACCGCCGCAAGGTTAAACTCAATGATTAATTTGACGGG	927
QY	896	GACCCCGCAAGCGGTGATTAATGGAATTAATTCGATGCAAGCGCAAAAACTTAACCTA	955
Db	928	GGCCCGCAAGCGGTGAGCATGTGGTTAATTCGATGCAAGCGCAAACTTAACCTA	987
QY	956	CCCTTGACATGTAGCGAAATTTCTAGAGAT-AGATTAGTCTTCGGGAAACGCTAACAG	1014
Db	988	GCTTTGACATCAAGGAAGTTTTCAGAGATGAGAAATGTGCTTGGGGAAACGTGACAG	1047
QY	1015	GTGCTGACATGCGTGTGCTGAGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGC	1074
Db	1048	GTGCTGACATGCGTGTGCTGAGTGTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGC	1107
QY	1075	GCAACCTTTGATTAATTTGCAATC-ATTGTTGGGCACTTAAATGAGATGCGCGGGA	1133
Db	1108	GCAACCTTTATCTTTGTTGCAACGCGTCCGGCGGAACTCAAGAGACATGCGCAATGA	1167
QY	1134	CAAAACCGGAAAGGTGGGAGTAGAGTCAATCTCCATAGCGCTTAATGGGTAGGGCTTCA	1193
Db	1168	TAAATCTGAGGAAGGTGGGAGTAGAGTCAATCTCATATGAGCGCTTATGACACAGGGCTACA	1227

QY	1194	CACGTAATACATGGCGGTACAGAGGGTTCGAACCCGCGAGGGGGAGTTAATCTAGA	1253
QY	1228	CACGGCTACATGGCGCATCAGAGAGAGAGGACCTTCGAGAGCAAGCGGACTATTA	1287
QY	1254	AAGCGCGTCGTAGTCCGAGTCGGAATCTGCAATCTGACTCCGTGAAGTCGAAATCGCTAG	1313
Db	1288	AAGTGCCTGTAGTCCGAGTTGGAATCTGCAACTCGACTCGATGAAAGTCGAAATCGCTAG	1347
QY	1314	TAATCGCGGATCAGCATATCGCGGGTGGAATACGTTCCGGGCTTSTTACACACCGGCCCTC	1373
Db	1348	TAATGTGATCAGAAATGCCACAGGTGAATACGTTCCGGGCTTGTACACACCGGCCCTC	1407
QY	1374	ACACCATGGAGTGGGTTTACACAGACGAGTAATCTTAAACCTGAAGAGGGCGCTTGCC	1433
Db	1408	ACACCATGGAGTGGGTTTCAAAAAGATAGTACTTAAACCTTGGGAGGGCGGCTTAAC	1467
QY	1434	AACGTGATTCATGATCGGGGTG	1457
Db	1468	ACTTGTGATTCATGATCGGGGTG	1491

```

RESULT 27
US-11-198-794-158
/ Sequence 158, Application US/11198794
/ Publication No. US20060035257A1
/ GENERAL INFORMATION:
/ APPLICANT: BROW, MARY ANN D.
/ APPLICANT: LYAMICHEV, VICTOR I.
/ APPLICANT: OLIVE, DAVID M.
/ TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
/ TITLE OF INVENTION: PATHOGENS
/ NUMBER OF SEQUENCES: 165
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: MEDLEN & CARROLL
/ STREET: 220 MONTGOMERY STREET, SUITE 2200
/ CITY: SAN FRANCISCO
/ STATE: CALIFORNIA
/ COUNTRY: UNITED STATES OF AMERICA
/ ZIP: 94104
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/11/198,794
/ FILING DATE: 05-Aug-2005
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/09/941,193
/ FILING DATE: 28-Aug-2001
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CARROLL, PETER G.
/ REGISTRATION NUMBER: 32,837
/ REFERENCE/DOCKET NUMBER: FORS-01756
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 705-8410
/ TELEFAX: (415) 397-8338
/ INFORMATION FOR SEQ ID NO: 158:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1542 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-11-198-794-158

```

Query Match	68.4%	Score 996.8;	DB 18;	Length 1542;
Best Local Similarity	82.3%	Pred. No. 2,1e-299;		
Matches 1205;	Conservative	0;	Mismatches 252;	Indels 7;
				Gaps 5;

Db	28	ATTGAAAGCTGGCGGCGAGCTTAACACATGCAAGTGCMAACGGTAACGTAACGAAAGAACTTGC	87
Oy	59	ATC - TGGTGGCAGATGGCGGAGCGGGTGAATTAATGCAATCGAAACGATCCAGAAAGAGGG	116
Db	88	TTCTTTCTCAACAGTAGTGGCGGAGCGGGTGAATTAATGCTGGGAAACATGCCCTGATGAGGGG	147
Oy	117	GGTAACGCATCGAAGATGTGCTAATACCGCATTAATCTTAAGAGAAAGACGGGATC	176
Db	148	GATAACTACTGAAACCGGTAGCTAATATCCCATTAATCCATCCCAAGACCAAAAGAGGGGACC	207
Oy	177	GAAAGACCTTGCCGCTTTTGGAGGGGGCGGATGTGATTTAGTCTAGTTGGTGGGGTTAAAGC	236
Db	208	TTGCGGCTCTTTCGCAATCGAATGTGCCCAATGGGAATTAAGTAACTAGTAGTGGGTACGCC	267
Oy	237	CTACCAAGGCGACGATCAGTATGTGCTGTAGAGAGACGACGACGACCACTGGGACTGAGA	296
Db	268	TCACCTAGGCGACGATCCCTAGCTGGCTGTAGAGAAATGACCAAGCACACTGGAACTGAGA	327
Oy	297	CACGGCCCAAGCTCTTACGCGAGGCGACAGTGGGGAAATTTTGCACATGGGCGCAAGCT	356
Db	328	CACGGTCAAGACTCTACGGGAGCGACAGATGGGGAAATATTGACAAATGGGCCCAAGCT	387
Oy	357	GATTCACCAATGCCGCGTGAAGTAAGAAAGCCCTTGCGGTTGTAAAGCTCTTTCAGTGCAG	416
Db	388	GATCAGCCCAATGCCGCGTGAAGTAAGAAAGCCCTTGCGGTTGTAAAGTCTTTCAAGCGGG	447
Oy	417	AAGAAAGGTTACGGTAAATTAATCGTGAACCCATGACCGTATTCGACAGAAAGCACCGGC	476
Db	448	AGAAAGGAGTAAGATTAAATTAATCTTCTCTCATTTAGCCTTACCGCAAGAAAGCACCGGC	507
Oy	477	TAACTACGTCGCACGACCGCGGTAATACGTAGGGTCMAAGCTTAATTCGAAATTAATCGG	536
Db	508	TAACTCCGTCGACGACCGCGGTAATACGTAGGGTCMAAGCTTAATTCGAAATTAATCGG	567
Oy	537	GCGTAAAGGTCGCGCAGCGCGCTTGTAAATGACATGTAAGTAATCCCGGGCTTAACTCGG	596
Db	568	GCGTAAAGGTCGCGCAGCGCGCTTGTAAATGACATGTAAGTAATCCCGGGCTTAACTCGG	627
Oy	597	GAATTCGCTTGAATCTACAAAGCTAGATGTGCGACAGGAGGTGGAATTCATGTGTA	656
Db	628	GAATTCGCTCTGATTAATCTGCGCAAGCTTGAAGTCTGTAAGAGGGGGTAGAATTCAGAGTGA	687
Oy	657	GCAGTGAAGTCGTAGAGATATGGAAGAATCGATGTCGACGAGCGACTTCTGGGTTAA	716
Db	688	GCGGTGAAGTCGTAGAGATATGGAAGAATCCGCTGTCGAAAGCGCGCCCTCTGGAAGAA	747
Oy	717	CACGTACGCTCATCGAAGCGTGGGGAGCAACAGAAATTAAATACCTCGTAGTCCA	776
Db	748	GACTACACCTCAGGTCGAAAGCGTGGGGAGCAACAGAAATTAAATACCTCGTAGTCCA	807
Oy	777	CGCCCTAAACATGTCAACTAGTTGTTGGGCTTAAATTAAGGCTTGG - TAAAGAACTAACG	835
Db	808	CGCCCTAAACATGTGTCAACTGGAAGTTGTGCCCTTAGAGGTGTGGCTTCGGAAGCTAACG	867
Oy	836	CGTGAAGTTGACCGCTGCGGGAGTACGCTGCGCAAGATTTAAATCTCAAGAAATTGACGGG	895
Db	868	CGTTAAATGTCGACCGCTGCGGGAGTACGCGCCGCAAGTTAAATCTCAATGATTTGACCGG	927
Oy	896	GACCCGCGCAAGCGGTGATTAATGTGGAATTAATCGATGCAACCGGAAAACTTAACTTA	955
Db	928	GCGCCGCGCAAGCGGTGAGCATGTGTTAATTCGAGTCAACCGGAAACCTTAACCTG	987
Oy	956	CCCTTGAATGAGGAAATTTTCTAGAGAT - AGATTAGTGCCTTGGGGAACGCTAACAG	1014
Db	988	GCTTTGAATTCACGGAAGTTTTCAGAGATGGAATGTGCTTGGGGAACGCTAGACAG	1047
Oy	1015	GTGCTGACATGCTGTGCTCAGCTGCTGTGCTGAGATGTTGGGTTAATGTCCTCGCAACGAGC	1074
Db	1048	GTGCTGACATGCTGTGCTCAGCTGCTGTGCTGAGATGTTGGGTTAATGTCCTCGCAACGAGC	1107
Oy	1075	GCAACCTTTCATTAATTTGCTATC - ATTGGTTGGGCACTTATATGACATGCGCGGGA	1133

Db	1108	GCAACCCCTTATCCTTTGTGGCAGCGGTCCGGCCGGAACTCAAGAAGACTGCGCAGTGA	1167
Qy	1134	CAAAACCGGAGGAAGGTGGGATGACGTCAAGTCCTCATGSCCTTTATGGTAGGGCTTCA	1193
Db	1168	TAAACTGGAGGAAGTGGGGATGACGTCAAGTCATCATGAGCCCTTACAGACGAGGGCTTACA	1227
Qy	1194	CACGTAATACAAATGGCCGCTACAGAGGTTGCCAAACCCGCGAAGGGGAGACTAACTCAGA	1253
Db	1228	CACGTGCTACATGGCCGCTACAAAGAAACGACCTCGCGAAGCAAGCCGACCTCATTA	1287
Qy	1254	AAAGGCGTCGTAAGTCCGGATCGGAATCGGAATCGAATCGACTCGGTGAAGTCGAAATCGCTAG	1313
Db	1288	AAAGTCCGTCGTAAGTCCGGATTCGGAATCTGCAACTCGATCTCCATGAAGTCGAAATCGCTAG	1347
Qy	1314	TAAATCGCGGAATCAGCATGTCCGCGTGAAATACGTTCCCGGGCTCTTGTACACACCGCCGCTC	1373
Db	1348	TAAATCGTGGAATCAAGATGCCAAGGTGAATACGTTCCGGGGCTTGTAACAACCGCCGCTC	1407
Qy	1374	ACACCAATGGAGTGGGTTCAACAGAGCAAGTAGTCTAAACCGTAAGAGGGCGCTTGCC	1433
Db	1408	ACACCAATGGAGTGGGTTGCAAAAGAAATAGTAGTAACTTCGGAGGGCGCCCTTAAC	1467
Qy	1434	ACGGTGAGATTCATGACTGGGGGTG	1457
Db	1468	ACTTTGATTCATGACTGGGGGTG	1491

```

1      RESULT 28
2      US-11-198-657-158
3      ; Sequence 158, Application US/11198657
4      ; Publication No. US20060040299A1
5      ;
6      ; GENERAL INFORMATION:
7      ; APPLICANT: BROW, MARY ANN D.
8      ; APPLICANT: LYAMICHEV, VICTOR I.
9      ; APPLICANT: OLIVE, DAVID M.
10     ; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
11     ; TITLE OF INVENTION: PATHOGENS
12     ; NUMBER OF SEQUENCES: 165
13     ; CORRESPONDENCE ADDRESSES:
14     ; ADDRESSEE: MEDLEN & CARROLL
15     ; STREET: 220 MONTGOMERY STREET, SUITE 2200
16     ; CITY: SAN FRANCISCO
17     ; STATE: CALIFORNIA
18     ; COUNTRY: UNITED STATES OF AMERICA
19     ; ZIP: 94104
20     ;
21     ; COMPUTER READABLE FORM:
22     ; MEDIUM TYPE: Floppy disk
23     ; COMPUTER: IBM PC compatible
24     ; OPERATING SYSTEM: PC-DOS/MS-DOS
25     ; SOFTWARE: PatentIn Release #1.0, Version #1.30
26     ;
27     ; CURRENT APPLICATION DATA:
28     ; APPLICATION NUMBER: US/11/198,657
29     ; FILING DATE: 05-Aug-2005
30     ; CLASSIFICATION:
31     ;
32     ; PRIOR APPLICATION DATA:
33     ; APPLICATION NUMBER: US/09/941,193
34     ; FILING DATE: 28-Aug-2001
35     ; ATTORNEY/AGENT INFORMATION:
36     ; NAME: CARROLL, PETER G.
37     ; REGISTRATION NUMBER: 32,837
38     ; REFERENCE/DOCKET NUMBER: FORS-01756
39     ; TELECOMMUNICATION INFORMATION:
40     ; TELEPHONE: (415) 705-8410
41     ; TELEFAX: (415) 397-8338
42     ;
43     ; INFORMATION FOR SEQ ID NO: 158:
44     ; SEQUENCE CHARACTERISTICS:
45     ; LENGTH: 1542 base pairs
46     ; TYPE: nucleic acid
47     ; STRANDEDNESS: double
48     ; TOPOLOGY: linear
49     ;
50     ; MOLECULE TYPE: DNA (genomic)
51     ;
52     ; US-11-198-657-158

```

Query Match	68.4%	Score 996.8;	DB 18;	Length 1542;
Best Local Similarity	82.3%	Pred. No. 2.1e-299;		
Matches 1205; Conservative	0;	Mismatches 252;	Indels 7;	Gaps 5;

OY	1	ATTGAAGCTGGCGCAGATGCTTTTACATGCAATCGAACCGGACGACG--GATGCTTGG	58
Db	28	ATTGAACGCTGGCGGACGCTTAACAATGCAATCGGAACGGTAAACGGAAGAACTTGG	87
OY	59	ATC--TGGTGGCAGATGGCGGACGGGTGAGTAATGCAATCGGAACGTAATCCGAAGAAGGGG	116
Db	88	TTCTTTGCTCGACGATGGCGGACGGGTGAGTAATGCTGGGGAACCTGCTGATGGAGGGG	147
OY	117	GGTAAACGATCGAAGAATGTCCTAATCCGGAATATACCTTAAGGAGCAACAGAGGAGATC	176
Db	148	GATATACTACTGGAAACGGTACTTAATACCGATTAACGTTCGAAGACCAAGAGGGGAGAC	207
OY	177	GAAGACCTTGGCGCTTTTGGAGCGGCGGATGTCTGATTAGTAGTTGGGTGTAAGGC	236
Db	208	TTGGGGGCTTGTGCGATCGGATGTGCCCAGATGGGATTAAGCTAGTAGGTGGGTAAAGGC	267
OY	237	CTAACCAAGCGACGATCAAGTATGTTGGTCTGAGAGAGCAACAAGCACATGGAACCTGGA	296
Db	268	TCACCTAGGCGACATCCCTAGCTGTGAGAGATGACCAAGCCACACTGAACTGAGA	327
OY	297	CACGAGCCAGACTCTTACGGGAGCAGCAGTGGGGAATTTTGGACAAATGGGCGCAAGCT	356
Db	328	CACGCTCCAGCTCTTAAGGGAGGACGAGTGGGGAATATTGCAAAATGGGCGCAAGCT	387
OY	357	GATCCAGCAATGCCGCGTGAAGAAAGGCTTCCGGTTGTAAAGCTCTTCACTGAG	416
Db	388	GATCAGACCAATGCCGCGTGTGAAGAAAGGCTTCCGGTTGTAAAGTCTTTCAGCGGGG	447
OY	417	AAGAAAGGTATCGGTAAATTAATGTGACCAATACGCTATCCGACAGAAAGACCGGC	476
Db	448	AGGAAAGGATGAATTAATTAATCTTTGCTACTGACGTTACCCGAGAAAGACACCGGC	507
OY	477	TAACTAGGTGCGACAGCCGCGTAAATACGTAGGGGTGCGACGTTAATCGAATTAATCTGG	536
Db	508	TAACTCGGTGCAACAAGCCGCGGTAAATACGGAAGGGTGCACGTTAATCGAATTAATCTGG	567
OY	537	GCGTAAAGGGTGGCGAGCGGCTTTGTAATCAATGTGAATATCCCGGGCTTAACTCTGG	596
Db	568	GCGTAAAGCGCACCGCGGTTGTTGAATCAATGTGAATATCCCGGGCTCAACTCTGG	627
OY	597	GAATTGGGTTGAACACTACAAAGCTAGGTGGCAGAGGGAGGTGAATTCATGTGTGA	656
Db	628	GAATCTGACTTGAAATCTGGCAAGCTTTGAGTCTCGTAAGGGGGGTGAATTCAGAGTGA	687
OY	657	GCACTGAATATCGTAGAGTATGGAAGCAATCATGCGAAGCAGGCTCTCTGGGTTAA	716
Db	688	GCGGTGAATATCGTAGAGTCTGGAGGAATACCGGTGGCGAAGCGGCGCCCTGGACGA	747
OY	717	CACTAGGCTCATGACGAAGAAGCTGGGGAGCAAAACAGATTAATATACCTGTGATGTCA	776
Db	748	GACTGACGCTCAGGTGCGAAAGCGTGGGGACAAACAGATTAATATCCTGTGATGTCA	807
OY	777	CGCCCTAAACGATGCAACTAGTATTGGGGCTTATTAGGCTTGG--TAACGAAGCTAAGC	835
Db	808	CGCGGTAAACGATGTCACTTGAGAGTTGTGCCCTTAGAGGCTGTGCTCCGGAAGCTAAG	867
OY	836	CGTGAAGTTGACGCGCTGGGGAGTACGGTGCGCAAGATTTAAACTCAAGAAATTTGACGG	895
Db	868	CGTTAAGTCAGCCGCTGGGGAGTACGCGCCCAAGGTTAAATCTCAAAATGAATTTGACGG	927
OY	896	GACCCGACAAAGCGGTGGAATTAATGTGAATTAATCGAGCAACGGGAACCTTAACCTGA	955
Db	928	GGCCCGCAAAAGCGGTGAGCATGTGTTAATTCAGTGAACCGGAAGAACCTTAACCTG	987
OY	956	CCCTTGAATGTACGAATTTTCTAGAGAT--AGAATTAGTGTCTTGGGAAACGCTAACAG	1014
Db	988	GTCCTTGACATCAAGGAATTTTCAAGATGAGATGTGCTTTGGGAAACGCTGAGCAG	1044
OY	1015	GTGCTGACAGGCTGTCTGACGCTGCTGTGTAAGTGTGGGTTAATGCTCCGCAACAGAC	1074

Db	1048	GTGCTGATGCTCTCGTCAAGCTGTGTTGGAAATGTGGTTAAAGTCCCGACAAGAC	1107
Qy	1075	GCAACCCCTTGTCATTTAATATGCGATC-ATTGSGTTGGGCACTTTAATGAGATCGCCGGTGA	1133
Db	1108	GCAACCCCTTATCTCTTTGTGTGCGACACGGGTCCGGCGGAACCTCAAAAGAGATCGCCAGTGA	1167
Qy	1134	CAAAACCGAGAGAGTGGGAGTGAAGTCGTCAAGTCTCATGAGCCCTTTATGGTAGGGCTTCA	1193
Db	1168	TAAATCTGAGAGAAAGTGGGAGTGAAGTCATAGTCAATAGCCCTTACAGCCAGGGCTTACA	1227
Qy	1194	CACGTAATATCATGGCCGCGTACAGAGGTTGCCAACCCGCGAGGGGAGAGCTAAATCTAGA	1253
Db	1228	CACGTGTCTACATAGCGCGCATACAAAGAAACGACCTCGCGAGAGCAAGCCGACCTCAATA	1287
Qy	1254	AAGGCGCGTCGATGCCGATCGAGATCGAACTCGAATCTCGTGAAGTCGGAATCGCTAG	1313
Db	1288	AAGTACCGTCGATGCTCCGATTTGAGATCTGCAACTCGAATCCATGAAAGTCGAAATCGCTAG	1347
Qy	1314	TAAATCGCGATCAGCATGTGCGCGGTAAATACGTTCCCGGGTCTTGTCACACCGCCCGCTC	1373
Db	1348	TAAATGTGAGATCAGAAATGCAACGGGAAATACGTTCCCGGGCTTGTCACACCGCCCGCTC	1407
Qy	1374	ACACCATGGAGTGGGTTTCAACAGAGCAGGTAGTCTAAACCTGTAAGAGGGGCGCTTGCC	1433
Db	1408	ACACCATGGAGTGGGTTTCAAAAGAAATAGGTAGTCTTAACCTTCGAGAGGGCGCTTATCC	1467
Qy	1434	ACGGTGAATTCATGACTGGGGTG	1457
Db	1468	ACTTGTGATTCATGACTGGGGTG	1491

## RESULT 29

```

US-10-831-286A/c
; Sequence 48668, Application US/10831286A
; Publication No. US20060046246A1
;
; GENERAL INFORMATION:
;
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELLIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
;
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
;
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; PRIOR FILING DATE: 2003-04-24
;
; NUMBER OF SEQ ID NOS: 48788
;
; SOFTWARE: PatentIn version 3.2
;
; SEQ ID NO 48668
;
; LENGTH: 1533
;
; TYPE: DNA
;
; ORGANISM: Cedecaea lapagei
;
US-10-831-286A-48668

```

Query Match 67.9%; Score 990; DB 11; Length 1533;

Best Local Similarity 82.0%; Pred. No. 2,7e-297;  
Matches 1202; Conservative 0; Mismatches 255; Indels 9; Gaps 5

QY 1 ATTGAACGCTGGCG-GCATGCTTTACATCGCAATCGAAGGCGACGCG-----GATGC 54  
Db 1511 ATTGAACGCTGTCGGCGCAGCGCTTAACAATGCAATGAATCGAAGCGGTATGCACGGGAAGCTTG 1452  
QY 55 TTGCATCTGGTGGCCAGTGGCGGACGGGTGAATTCATCGGAACGTATCCAGAAAGG 114  
Db 1451 CTCCTCGGTGACGAGCGCGCGAGCGGGTGAATTCCTGGGGATCTGCTCGATGAGG 1392  
QY 115 GGGGTAAACGATCGAAGAATGTGTAATACCGCATATACCTTAAGAGAGAAAGCAGCGGA 174  
Db 1391 GGGATTAACATCTGGAACCGGTACTAATACCGCATTAAGCTCCCAAGACCAAGAGGGGGA 13322  
QY 175 TCGAAGACCTTGCCTTTTGGAGCGGCCGATGTCATTAAGTAAGTTGGTGGGCTAAG 234

Db	1331	CGTTGCGGCGCTTGGCATCAGATGGAACCCAGATGGGATTAGCTAGTAGGTGGGGTAAACG	1272
Oy	235	GCTTACCAAGGCGGACGATCAGTAGTGGTCTTGAGAGAGCGACGAGCGACACTGGGACTGA	294
Db	1271	GCTCACCTAGGCGCGAGATCCCTAGCTGGTCTGAGAGGATGACGAGCGACACTGGAACCTGA	1212
Oy	295	GACAGCGGCCGACATCTCTAGGGGAGGACAGAGTGGGGAATTTTGGACAATGGGCGCAAGC	354
Db	1211	GACACGGTCCAGATCTCTACGGGAGGCGACAGTGGGGGAATTTGGACAATGGGCGCAAGC	1152
Oy	355	CTGATCCAGCAATCCCGCTGAGTGAAGAAAGGCGCTTCGGTTGTAAAGCTCTTTCAGTCG	414
Db	1151	CTGATGACGCAATCCCGCTGTATGAAAGAGCGCTTCGGTTGTAAAGTACTTTCAGCGCA	1092
Oy	415	AGAAAGAAAGGTTACGGTAAATATTCGTGACCCATGACGCTATCGACAGAAAGACCGC	474
Db	1091	GGAGAGAGGCAATTAAGGTTAATACTTAACGATTGACGTTACTCGCAGAAAGAACCGC	1032
Oy	475	GCTTAATCAGTCCGACAGACCGCGGTTAATGTAAGGTTGCAAGGCTTAATCCGAATTACT	534
Db	1031	GCTAATCTCCGTCAGAGACCGCGGTTAATCGAGGGTGTCAAGGCTTAATCCGAATTACT	972
Oy	535	GGGGGTAAAGGGTCCGACAGGCGCGCTTGTAAAGTCAGATGTGAATCCCGGGCTTAACT	594
Db	971	GGGGGTAAAGGCGACGACGCGGCTTGTGTAAAGTCGAGATGTGAATCCCGGGGCTCAACT	912
Oy	595	GGGAATTGCGCTTTGAAACTACAAAGCTAGAGTGTGGCAGAGGGAGTGGAAATTCATGTG	654
Db	911	GGGAATGTCATTTGAAACTGTGCAAGCTTGAAGTCTTGTAGAGGGGGGTAGAAATTCAGGTG	852
Oy	655	TAGCAGTGAATGGGTAGAGATATGAGAAACAATCGATGTGCGAAGCGCAAGCGCTCTGGGTT	714
Db	851	TAGCGGTGAATGGGTAGAGATCTGAGAGGAATACCGGTGGGAAAGCGCGCCCCCTGACACA	792
Oy	715	AAACCTGACGTCACATGACAGAAACGATGGGAGCAAAACAGATTAGATTAACCTGTGATGTC	774
Db	791	AAGATGACGTCACAGGTGCGAAACGATGGGAGCAAAACAGATTAGATTAACCTGTGATGTC	732
Oy	775	CACGCGCTTAAACGATGTCTCACTAGTGTGTGGGCGCTTATTAGGCTTGG-TAACGAAGCTTAA	833
Db	731	CACGCGGTAAACGATGTCTCACTTGTGGAAGTTGTGCGCTTGAAGGCGGTGCTTCGGAGCTTAA	672
Oy	834	CGCGTGAAGTTGACCGGCTGGGAGATACGCTGCGCAAGTTTAAACTCAAGAAATTTAGC	893
Db	671	CGCGTTAAGTGAACGCGCTGGGAGATACGCGCGCAAGTTTAAACTCAAGAAATTTAGC	612
Oy	894	GGGACCCGCGACACGCGGTGGAATTTATGTCGATTTAATTGCAATGCGCAAAACCTTACC	953
Db	611	GGGGCGCCGACAAAGCGGTGAGCAATGTGGTTAATTCGATGCAACGCGCAAGAACCTTACC	552
Oy	954	TACCCCTGAACATGTAAGGGAATTTTCTAGATAGATTAGTG-CTTCGGGAAACGCTTAAAC	1012
Db	551	TACTCTTGAACATCCAGAGAACTTTCGAGATAGATTGTGTCTTCGGGAACTGTGAGAC	492
Oy	1013	AGGTGCTCATGCGTGTGCTGACGTCGTGCTGAGATGTTGGTTAAGTCCCGCAACGA	1072
Db	491	AGGTGCTCATGCGTGTGCTGACGTCGTGCTGAGATGTTGGTTAAGTCCCGCAACGA	432
Oy	1073	GGGCAACCTTGTCAATTAATTGCCATC-ATTGTGTTGGGCACTTTAATGAACATGCCGGT	1133
Db	431	GGGCAACCTTATCTTGTGTGCGACGCGTTTCGCGCGGGAATCAAAAGAAATCCGCAAT	372
Oy	1133	GACAAACCGGAGAGAGTGGGGATGACGTCAGATCTCATGAGCCCTTATGGGTAGGCGCTT	1193
Db	371	GATTAACCTGAGAGAGTGGGGATGACGTCAGATCATATGAGCCCTTATGAGTAGGCGCTA	312
Oy	1192	CACACGTAAATGAATGCGCGCTTACAGAGGTTGGCCAAACCGCGAGAGGGAGCTAATCTCA	1251
Db	311	CACACGTGTCAATGCGCGCATACAAAGAAAGGACCTTCGCGAAGCAACGGAACCTCA	252
Oy	1252	GAAGGCGGTGTAGTCCGGAATCGGAGTCTGCAATTCGATCTCGTGAAGTCCGGAATCGCT	1311



```

Db      251 TAAAGTCGCTCGTAGTCGCGGATTGAGTCTGCAACTCGACTCGATGAGTGGAAATCGCT 192
Qy      1312 AGTAATCGCGGATCAGCATGTCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCG 1371
Db      191 AGTAATCGTAGATCAGAAATGCTACGGGTGAATACGTTCCCGGGCTTGTACACACCGCCG 132
Qy      1372 TCACACCATGAGGTGGGTTTTCACGAAAGCAGTAGTGTCTAAACGTTAAGGGGGCGCTTG 1431
Db      131 TCACACCATGAGGTGGGTTTTCACGAAAGCAGTAGTGTCTAAACGTTAAGGGGGCGCTTG 72
Qy      1432 CCACGGGAGATTGATGACTGGGGTG 1457
Db      71 CCACCTTGTGATTCAATGACTGGGGTG 46

RESULT 30
US-10-831-286A-48691
; Sequence 48691, Application US/10831286A
; Publication No. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174, .001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48691
; LENGTH: 1418
; TYPE: DNA
; ORGANISM: Morganella morganii sp. sblonii
US-10-831-286A-48691

Query Match      67.8%; Score 988; DB 11; Length 1418;
Best Local Similarity 83.2%; Pred. No. 1,1e-296;
Matches 1160; Conservative 0; Mismatches 230; Indels 4; Gaps 3;

Qy      68 CGAGTGGCGGACGCGGTGAGTAATGATCGGACGTATCCGAAAGAGGGGGGTAAACGATC 127
Db      1 CGAGCGGCGGACGCGGTGAGTAATGATGCGGATCTGCTATGGAGGGGATTAATTAATG 60
Qy      128 GAAAGATGTCTAATACCGGATATATCTCTAAGAGGAAAGCAGGGGATCGAAAGACCTTG 187
Db      61 GAAACGGTACCTAATATCCGATATATGTCACGGACCAAGCGGGGACCTTCGGGCTCG 120
Qy      188 CGCTTTTGACCGCCGATGTCTGATTAGCTAGTGTGGGTAAAGGCTTACCAAGCG 247
Db      121 CACCATCATGATGAACCATATGGATTAGCTTGTAGGTGAAGTAACGGCTCACTTAGGCG 180
Qy      248 ACAGTACAGTATGTGTCTGAGAGACGACCAAGCCACTGAGACTGAGACACGCGCCCA 307
Db      181 ACATCCCTAGTGTGTCTGAGAGATGATCAGCACACTGGGACTGAGACACGCGCCCA 240
Qy      308 CTCTTACGGAGGACGAGAGTGGGAAATTTTGAACAAATGGGCGCAAGCTTATCCAGCAAT 367
Db      241 CTCTTACGGAGGACGAGAGTGGGAAATTTTGAACAAATGGGCGCAAGCTTATCCAGCAAT 300
Qy      368 GCCGCGTAGTGAAGAGGCTTCGCGGTGTTAAAGCTCTTTCAGTCCGAGAAAGAAAGTT 427
Db      301 GCCGCGTAGTGAAGAGGCTTCGCGGTGTTAAAGCTCTTTCAGTCCGAGAAAGAAAGTT 360
Qy      428 ACAGTAATTAATCGTGAACCATGACGCTATCGACAGAAAGCAACGCGCTTAATCACTGTC 487
Db      361 AAGGTTAATAACCTTGGCAATTGACGTTACGACAGAAAGCAACGCGCTTAATCACTGTC 420
Qy      488 CAGCAGCGCGGTATATCGTAGGGTGCAGCGCTTAATCGGAATTAATCTGGGCGTAAAGGT 547

```

```

Db      421 CAGCAGCCGCGTAATACGAGAGGTGCAAGCGTTAATCGAATTAATGCGCGTAAAGCGC 480
Qy      548 GCGCAGCGCGCTTGTAAATGATGATAATATATATATATATATATATATATATATATATATAT 607
Db      481 AGCAGAGCGGTGATGATGATGATGATAATATATATATATATATATATATATATATATATATAT 540
Qy      608 GAAATCAACAAAGCTAGAGTGTGGCAGAGAGGTGAAATTTCCATGTGTAGCAGTAAATG 667
Db      541 GATTAATCGTACGTAAGTCTTGTAGAGGGGGTGAATTCATGTGTAGAGCGGTGAATG 600
Qy      668 CGTAGAGATATGAAGAACATGATGGCGAAGCAGCTCTCGGGTTAACTGACGCTC 727
Db      601 CGTAGAGATGTAAGAAATACCGGTGGCGAAGCGCGCCCTGGAACAAAGATGACGCTC 660
Qy      728 ATGACAGAAAGCTGTGGGAGCAAAAGATTTGATATCTTGATCTCCAGCGCCCTAAAG 787
Db      661 AGGTGCAAAAGCGTGGGAGCAAAAGATTTGATATCTTGATCTCCAGCGCTGTAACG 720
Qy      788 ATGTCAACTAGTTGTTGGCTTATTTAGCTTGG-TAAAGAGCTTAACGCGTGAAGTTGA 846
Db      721 ATGTGACTTTGAGAGTTGTGCTTGAAGCGGTGCTTCGAGAGCTTAACGCGTTAAGTGA 780
Qy      847 CCGCTGGGAGTACGCTGCAAGATTTAACTCAAAAGAAATGAACGGGACCCGACAA 906
Db      781 CCGCTGGGAGTACGCGCGCAAGGTTAAATCTCAAAATGAATGAACGGGAGCCGACAA 840
Qy      907 GCGGTGATATATGATGATTAATTCATGATGCAACGCAAAACCTTACCTTACCTTACATG 966
Db      841 GCGGTGAGCAATGTGGTTAATTCATGATGCAACGCAAAACCTTACCTTACCTTACATG 900
Qy      967 TAGCGAATTTCTAAGATAGATTAATG-TTCGCGAAAGCTTAACAGAGTCTGATG 1025
Db      901 CAGAGAACTTACAGAGATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 960
Qy      1026 CTGTCTCAGCTCGTGTCTGATGATGTTGGTTAAGTCCGCAAGCGCAACCTTGT 1085
Db      961 CTGTCTCAGCTCGTGTCTGATGATGTTGGTTAAGTCCGCAAGCGCAACCTTGT 1020
Qy      1086 CATTAATTCCTCATTTG--GTTGGGCACTTTAATGAGATCTGCGGTGACAAACCGGAG 1143
Db      1021 CTTTGTGTCAGCGCGGTATGCGGAACTCAAAAGAGATGCGCGGTATTAACCGGAG 1080
Qy      1144 GAAAGTGGGATGAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACAGTAAATG 1203
Db      1081 GAAAGTGGGATGAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACAGTAAATG 1140
Qy      1204 AATGGCGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGGCTG 1263
Db      1141 AATGGCGTATCAAAAGGAAAGCGAACCTGCGAAGGCAACGGAATCATAAAGTACGTG 1200
Qy      1264 TAGTCCGATCGGATCTGCAACTGACTCTCGTGAAGTGGAAATCGCTAGTAATCGCGGA 1323
Db      1201 TAGTCCGATCGGATCTGCAACTGACTCTCGTGAAGTGGAAATCGCTAGTAATCGCTAG 1260
Qy      1324 TCAGCATGTGGGGTAATACGTTCCGGGTCTTGTACACACGCGCGCTACACATAGG 1383
Db      1261 TCAGCATGTGGGGTAATACGTTCCGGGTCTTGTACACACGCGCGCTACACATAGG 1320
Qy      1384 AGTGGGTTTCAACAGAAAGAGTATCTAACCGTAAAGAGGCGCTTGCACGCTGAGAT 1443
Db      1321 AGTGGGTTTCAAAAGAAAGTATAGTAACTTCCGAGGGCGCTTACCATTTGTAT 1380
Qy      1444 TCATGACTGGGGTG 1457
Db      1381 TCATGACTGGGGTG 1394

RESULT 31
US-10-831-286A-48692
; Sequence 48692, Application US/10831286A
; Publication No. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG

```



```

; APPLICANT: CHATELLIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; PRIOR FILING DATE: 2003-04-24
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 48692
; LENGTH: 1418
; TYPE: DNA
; ORGANISM: Morganella morganii sp. silbonii
US-10-831-286A-48692

```

```

Query Match      67.8%; Score 988; DB 11; Length 1418;
Best Local Similarity 83.2%; Pred. No. 1.1e-296;
Matches 1160; Conservative 0; Mismatches 230; Indels 4; Gaps 3;

```

```

QY 68 CGAGTGGCGGACGGGTGATGATGATCGGACCTATCCAGAGAGGGGGGTACGCGATC 127
DB 1 CGAGCGCGGACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
QY 128 GAAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAGACCTTG 187
DB 61 GAAACGGTAGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAGACCTTG 120
QY 188 CGCTTTTGAGCGCGCCGATGTCTGATTAGCTAGTGTGAGGATTAAGGCTTACCAAGCG 247
DB 121 CACATCAGATGAAACCCATATGGATTAGCTTGTAGGTAGAGTAAAGCTTACCTAAGCG 180
QY 248 ACGATCAGTGTGTGTCTGAGAGAGACGACGACCTGAGGATGAGAGAGAGAGAGAGAG 307
DB 181 ACGATCCCTGAGTGTGTCTGAGAGAGATGATGACGACCTGAGGATGAGAGAGAGAGAG 240
QY 308 CTCTTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 367
DB 241 CTCTTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 368 GCCCGGTGAGTGAAGAGGCGCTTGGGTTGTAAGCTCTTCACTGTCAGAGAGAGAGGTT 427
DB 301 GCCCGGTGATGAAGAGGCGCTTGGGTTGTAAGCTCTTCACTGTCAGAGAGAGAGGTT 360
QY 428 ACGGTAAATATCTGTGACCCATGACGCTATGACAGAGAGAGAGAGAGAGAGAGAGAG 487
DB 361 AAGGTAAATATCTGTGCAATTTGACGTTACGACAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 488 CAGAGCGCGGCTAATCGTGGGTGCAAGGCTTAACTGGAATTAAGTGGGCTTAAAGGAT 547
DB 421 CAGAGCGCGGCTAATCGGAGGTGCAAGGCTTAACTGGAATTAAGTGGGCTTAAAGGAT 480
QY 548 GCGCAGCGCGCTTGTAGTCAATGTGAATCCCGGGCTTAACTGGAATTAAGTGGGCTT 607
DB 481 ACGCAGCGCGCTTGTAGTCAATGTGAATCCCGGGCTTAACTGGAATTAAGTGGGCTT 540
QY 608 GAAACTACAAAGCTAGAGTGTGCGAGAGAGAGTGAATTCATGTGTAGAGAGAGAGAG 667
DB 541 GATCTGCTCAGCTAGAGTCTTGTAGAGAGAGAGTGAATTCATGTGTAGAGAGAGAGAG 600
QY 668 CGTGAAGATATGAGAGAGAGATGAGTGGAGAGGAGGCTCTGGGCTTAACTGAGCGCTC 727
DB 601 CGTGAAGATATGAGAGAGATGAGTGGAGAGGAGGCTCTGGGCTTAACTGAGCGCTC 660
QY 728 ATGACGAAAGCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 787
DB 661 AGGTGCGGAAAGCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 788 ATGTCAACTAGTGTGTGGGCTTATTAAGCTTGG- TAAAGAGAGCTTAAGCGGTGAAGTTGA 846
DB 721 ATGTGCACTTGAAGGTGTGTGCTTGAAGGCGTGGCTTCCGAGGCTTAAGCGGTGAAGTTGA 780

```

```

QY 847 CCGCCTGGGAGTACGCTGCGCAAGATTAACTCAAGGATTTGACGGGGAACCGGCAAA 906
DB 781 CCGCCTGGGAGTACGCGCAAGATTAACTCAATATTAATTGACGGGGGCCGCAAA 840
QY 907 GCGGTGATTAATGATTAATTCAGTACGCGCAAGAACTTACCTTACCTTGAATG 966
DB 841 GCGGTGATTAATGATTAATTCAGTACGCGCAAGAACTTACCTTACCTTGAATG 900
QY 967 TAGCGAATTTTCTAGAGATGATTAAGT- CTTCGGGAGAGCTTAACAGAGTGTGATG 1025
DB 901 CAGAGAACTTACAGAGATGCTTGTGCTTCCGGAACCTTGAAGCAGGTGCTGATG 960
QY 1026 CTGTCTCAGCTCTGCTGCTGAGAGTGGGTTAAGTCCCGCAAGAGGCAACCTTGT 1085
DB 961 CTGTCTCAGCTCTGCTGCTGAGAGTGGGTTAAGTCCCGCAAGAGGCAACCTTGT 1020
QY 1086 CATTAAATGCAATCTTGTG- GTTGGGCACTTAAATGAGATGCTCCGTAACAAACCGAG 1143
DB 1021 CTTTGTGTCAGCGGCTGATGCGGGAACTCAAGAGAGATGCGGTGATTAACCGAG 1080
QY 1144 GAAAGTGGGAGTACGTCAGTCTGATGAGGCTTATGAGGAGCTTACACGTAATAC 1203
DB 1081 GAAAGTGGGAGTACGTCAGTCTGATGAGGCTTATGAGGAGCTTACACGTAATAC 1140
QY 1204 AATGCGCGTACAGAGAGGTTGCAACCGCGAGGGGAGCTATCTCAAGAAACGGGTG 1263
DB 1141 AATGCGCGTATCAAGAGAGGCAACCTGCGAAGGCAAGGAGAACTCAATTAAGTAC 1200
QY 1264 TACTCGGATCGAGTCTGCACTCGGAGAGTGGGATCGGATTAATTCGCGGA 1323
DB 1201 TACTCGGATCGAGTCTGCACTCGGAGAGTGGGATCGGATTAATTCGCGGA 1260
QY 1324 TCAGCATGTGCGGTAATACGTTCCCGGCTTGTACACACCGCCGTCACACATGG 1383
DB 1261 TCAGCATGTGCGGTAATACGTTCCCGGCTTGTACACACCGCCGTCACACATGG 1320
QY 1384 AGTGGGTTTCAACAGAGAGAGTGTATCCGTAAGAGAGGAGGCTTGCACGCTAGAT 1443
DB 1321 AGTGGGTTTCAACAGAGAGAGTGTATCCGTAAGAGAGGAGGCTTGCACGCTAGAT 1380
QY 1444 TCATGACTGGGGTG 1457
DB 1381 TCATGACTGGGGTG 1394

```

## RESULT 32

US-10-831-286A-48667/C

Sequence 48667, Application US/10831286A

Publication No. US20060046246A1

GENERAL INFORMATION:

APPLICANT: ZENG, QIANDONG

APPLICANT: CHATELLIER, SONIA

APPLICANT: MOIR, DONALD T.

APPLICANT: LACROIX, BRUNA

APPLICANT: CHILDRESS, DARRELL

TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES

FILE REFERENCE: 032796-174.001

CURRENT APPLICATION NUMBER: US/10/831,286A

PRIOR FILING DATE: 2004-04-26

PRIOR APPLICATION NUMBER: 60/464,955

NUMBER OF SEQ ID NOS: 48788

SOFTWARE: PatentIn version 3.2

SEQ ID NO: 48667

LENGTH: 1436

TYPE: DNA

ORGANISM: Cedecia lapagel

US-10-831-286A-48667

```

Query Match      67.3%; Score 981.2; DB 11; Length 1436;
Best Local Similarity 82.8%; Pred. No. 1.5e-294;
Matches 1157; Conservative 0; Mismatches 238; Indels 3; Gaps 3;

```

QY 63 GGTGGGAGTGGCGGACGGGTGATGATCGAAAGCTATCCAGAAAGAGGGGGTAAAC 122  
DB 1431 GGTGACGAGGCGCGGACGGGTGAGTAATGTCGGGGATCTGCTGATGAGGGGGATTAAC 1372  
QY 123 GCATCGAAAGATGTCGTAATACCGCATATATCTTAAAGAGAAAGAGGGGATCGAAAG 182  
DB 1371 TACTGAAACGAGTAGCTAATACCGCATATATCTGCAAGACCAAGAGGGGATCTTCCGG 1312  
QY 183 CTTGGCGCTTTTGGAGCGGCGGATGTCGATTAAGTATGTTGGGGTAAAGGCTTAACA 242  
DB 1311 CCTCTTGCCATCAGATGAACCCAGATGGGATTAAGTATGTTGGGGTAAAGGCTTAACCT 1252  
QY 243 AGCGCAGATCAGTATGTTGTTGAGAGCAGACAGCCACATGGGACTGAGACAGCGC 302  
DB 1251 AGCGCAGATCCTTACCTGTTGTCGAGAGTGAACAGCCACATCGAACTGAGACAGCGT 1192  
QY 303 CCGAGATCTTACCGGAGGAGCAGTGGGGAAATTTGGAGCAATGGGGGCAAGCCTGATCCA 362  
DB 1191 CCGAGATCTTACCGGAGGAGCAGTGGGGAAATTTGCAATGGGGGCAAGCCTGATCCA 1132  
QY 363 GCAATGCCGCTGATGAGAAAGGCTTCGGGTTGTAAGCTTTCACTGAGAAAGAA 422  
DB 1131 GCCATGCCGCTGATGAGAAAGGCTTCGGGTTGTAAGTACTTTCAGGAGAGAGAAAG 1072  
QY 423 AGGTTACGTTAAATATCGTACCCATGACGCTATGACAGAAAGAAAGCAGCGCTAACTA 482  
DB 1071 GCGATTAAGTTAAATCTTGTGATGATGAGCTTACCTGCAAGAAAGCAGCGCTAACTC 1012  
QY 483 CGTGCAGAGCGCGGCTATATGATGAGGCTTAAAGGCTTAAATGGAATTACTGGGGGTAA 542  
DB 1011 CGTGCAGAGCGCGGCTATATGAGAGGCTTAAAGGCTTAAATGGAATTACTGGGGGTAA 952  
QY 543 AGGGTCCGACGAGCGGCTTGTAAATGATGATGAAATCCCGGGCTTAACTGGGAAATG 602  
DB 951 AGGGCAGCAGGCGGCTTGTAAATGATGATGAAATCCCGGGCTTAACTGGGAAATG 892  
QY 603 CGTTGAAACTCAAAAGCTAGATGATGAGGAGGAGGAGTGAATTCATGTTGACAGT 662  
DB 891 CATTCGAACTGCGCAAGCTTGAATCTTGTAGAGGGGGGTGAATTCAGAGTGAAGCGGT 832  
QY 663 AAATGGTGAATATGAGAAAGATGATGAGGAGGAGGAGGCTTCGGGGTAACTGA 722  
DB 831 AAATGGTGAATATGAGAAAGATGATGAGGAGGAGGAGGCTTCGGGGTAACTGA 772  
QY 723 CGCTACGACGAAAGGCTGGGAGCAACAGGATTAAGATACCTGTAGTCCAGCGCT 782  
DB 771 CGCTACGACGAAAGGCTGGGAGCAACAGGATTAAGATACCTGTAGTCCAGCGCT 712  
QY 783 AAACGATGCACTAGTGTGGGCTTATTAAGGCTTGG-TAAAGAACTTAAGCGCTGAA 841  
DB 711 AAACGATGCACTAGTGTGGGCTTATTAAGGCTTGG-TAAAGAACTTAAGCGCTGAA 652  
QY 842 GTTGACCGGCTGGGAGTACGGTCCGAAAGATTAATCTCAAGAAATGAAGGGGAGCCG 901  
DB 651 GTTGACCGGCTGGGAGTACGGTCCGAAAGATTAATCTCAAGAAATGAAGGGGAGCCG 592  
QY 902 CACAAGCGGTGATTAATGATTAATTCGATGCAACGCGAAAAACCTTACCTTACCCTTG 961  
DB 591 CACAAGCGGTGATTAATGATTAATTCGATGCAACGCGAAAGCTTACCTTACCCTTG 532  
QY 962 ACATGTAAGCAATTTCTAGAGATAGTATGTC-CTTGGGAAAGCTTAACAGAGTGTG 1020  
DB 531 ACATGTAAGCAATTTCTAGAGATAGTATGTC-CTTGGGAAAGCTTGAACAGAGTGTG 472  
QY 1021 CAGTGGCTGTCAGCTCGTGTGTGATGATGTTGGGTTAAGTCCCGCAAGAGCGCAAC 1080  
DB 471 CAGTGGCTGTCAGCTCGTGTGTGATGATGTTGGGTTAAGTCCCGCAAGAGCGCAAC 412  
QY 1081 CTTGTCAATTAATGGCATC-ATTGTTGGGCACTTAAATGAGACTGCGCGGTGCAAAAC 1139  
DB 411 CTTATCTTTGTTGCCAGCGGCTTCGCGGGAACTCAAAAGAGACTGCGGTGATTAAC 352

QY 1140 GAGAGAGGTGGGAGTGAAGTCAAGTCCATGAGCCCTTAATGGTGAAGGCTTACAGCTA 1199  
DB 351 GAGAGAGGTGGGAGTGAAGTCAAGTCCATGAGCCCTTAATGGTGAAGGCTTACAGCTG 292  
QY 1200 ATACATAGCGCGGTACAGAGGGTTGCCAACCCGAGGGGGAGCTTAATCTCAGAAAGCGC 1259  
DB 291 CTACATAGCGCGGTACAGAGGGTTGCCAACCCGAGGGGGAGCTTAATCTCAGAAAGCGC 232  
QY 1260 GTTCGATGTCGAGATGAGAGTCTGCAACTGATCCGTTGAAGTCCGGAATTCGATTAATG 1319  
DB 231 GTTCGATGTCGAGATGAGAGTCTGCAACTGATCCGTTGAAGTCCGGAATTCGATTAATG 172  
QY 1320 CGGATCAGATGTCGGGGTGAATGATGTTCCGGGCTTGTACACACCGCCGCTACACCA 1379  
DB 171 TGATTCAGATCTACGAGTGAATGATGTTCCGGGCTTGTACACACCGCCGCTACACCA 112  
QY 1380 TGGAGTGGGTTGCAACAGAACAGTATGCTTAACGTAAGAGGAGCGCTTCCAGGTG 1439  
DB 111 TGGAGTGGGTTGCAAAAGAGTAAAGTAACTTAACCTTCGGAGGAGCGCTTACACTTTG 52  
QY 1440 AGATTCACTGCTGGGTG 1457  
DB 51 TGATTCACTGCTGGGTG 34

RESULT 33  
US-10-831-286A-48672  
; Sequence 48672, Application US/10831286A  
; General Information:  
; Publication No. US20060046246A1  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILDRESS, DARRELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; CURRENT FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; PRIOR FILING DATE: 2003-04-24  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 48672  
; LENGTH: 1481  
; TYPE: DNA  
; ORGANISM: *Citrobacter Youngae*  
US-10-831-286A-48672

Query Match 67.2%; Score 979.8; DB 11; Length 1481;  
Best Local Similarity 81.7%; Pred. No. 4,1e-294;  
Matches 1181; Conservative 0; Mismatches 257; Indels 7; Gaps 4;  
QY 20 CTTTACACATGCAAGTCGAAACGGCAGCAGATG----CTTGCATCTGTGTGCGAGTGGC 75  
DB 1 CTTACACATGCAAGTCGAAACGGTGAACAGAGAGCTTGTCTCTTGGGTGACGAGTGGC 60  
QY 76 GGAAGGGTGAATGATCGAAGCTATCCAGAAAGGGGGGTAAAGCATGAAAGATG 135  
DB 61 GGAAGGGTGAATGATGTCGGAAGAACTGCCATGAGAGGGGATTAATCTGGAAGCGT 120  
QY 136 TGCTAATACCGGATATACCTTAAGAGAGAAACAAGGGGATTCGAAAGACCTTGCGCTTTG 195  
DB 121 AGCTAATACCGGATTAACGTCGGAAGCAAGAGGGGAACTTTCGGGCTTTCGCAATG 180  
QY 196 GAGCGGCGCATGCTGATTAAGTATGTTGGGTAAAGGCTTACCAAGCGAGATCAG 255  
DB 181 GATGTCCCAAGATGGATTAAGTATGATGAGTGGGTAAAGGCTTACCAAGCGATGCC 240  
QY 256 TAGTGTCTGAGAGAGCAACCAAGCAACTGGGACTGAGACAGGCGCCAGACTCTTAG 315  
DB 241 TAGTGTCTGAGAGAGATGACCAAGCAACTGGAACCTGAGACAGGCTTCAAGACTCTTAG 300

QY	316	GAAGGCAAGCTGGGGAAATTTTGGACAATGGGGCGAAGCTGATTCAGAAATGCGCGTG	375
Db	301	GGAGGCGAGATGGGGAAATTTGGACAATGGGGCGAAGCTGATGCAAGCCATGCGCGTG	360
QY	376	AGTGAAGAAGGCCCTGGGGTTGTAAAGCTCTTTCAGTCGAGAAAGAAAGGTTACGGTAA	435
Db	361	TATTAAGAAGGCTTCGGGTTGTAAAGTACTTTTCAGGAGAGAGAAAGGCTTGAGTTAA	420
QY	436	TAATCGTAGCCCATGACGGTATCGACAGAAAGCACCGGCTAACTACGTCCAGCAGCC	495
Db	421	TAACTTAAAGCAATTTGACGTTACTCGAGAAAGCAAGCACCGGCTAACTCCGGCAGCAGCC	480
QY	496	GCGGTAAATACGTAGGGGTGCAGACCTTAAATGGAATTACTGGGGGTAAAGGATCGGAGGC	555
Db	481	GCGGTAAATACGGAAGGGGTGCAGACCTTAAATGGAATTACTGGGGGTAAAGGAGCGAGGC	540
QY	556	GGCTTTGAATGTCAGATGTGAATACCCCGGGCTTAACTGGGAATTGCGTTGAAACTAC	615
Db	541	GGTGTTCMAATCGGAATGTGAATACCCCGGGCTTAACTGGGAATTGCGTTGAAACTAC	600
QY	616	AAAGCTAGAGTGTGGCAGAGGGAGGTGAATTCATGTGTAGCAGTGAATGCGTGAAGA	675
Db	601	CAGGTAAGTCTGTAGAGGGGGGTGAATTCAGGTGTAGCGGTGAATTCGCTGAAGA	660
QY	676	TATGGAAGACATGATGTGGCGAAGGCGAGCTCTCGGTTAACTGAACGCTCATGACCA	735
Db	661	TCTGGAGGAATACCGGTGGCGAAGCGGCCCCCTGGAACAAGATGAACGCTCAGGTGCA	720
QY	736	AAAGGTGGGGAGCAACAGAGTTTGAATACCTGTGTAGTCCAGCGCCCTTAAACGATGTCAAC	795
Db	721	AAAGGTGGGGAGCAACAGATTTGAATACCTGTGTAGTCCAGCGCCCTTAAACGATGTCAAC	780
QY	796	TAGTTGTTGGGACCTTAATTAGGCTTGG-TAACGAAGCTTAACGGGTGAAGTTGACCGGCTGG	854
Db	781	TTTGAGAGTGTGGCCCTTGAAGGCGTGGTTCGGAGCTTAACGGCTTAAAGTGCACGCGCTGG	840
QY	855	GGAGTACGGTGCAGAGATTAAATCTCAAGGAATTGACGGGAGCCGCCACAAAGCGGTGGA	914
Db	841	GGAGTACGGGCGCGCAAGGTTAAATCTCAATTAATTAACGGGGGCGGCCACAAGCGGTGGA	900
QY	915	TTATGTGATTAATTCGATGCAACGGCGAAAAACCTTACCTTACCTTGAACATGTAGCAAT	974
Db	901	GCATGTGTTTAATTCGATGCAACGGCGAAAAACCTTACCTTACCTTGAACATCCAGAGAAC	960
QY	975	TTTTCAGAGATAGATTAGTG-CTTCGGGAAACGCTTAACAAGGTCTGCAATGGCTGTGCTC	1033
Db	961	TTTTCAGAGATGCAATTGGTGTGCTTCGGGAACTCTGAAGACAGGTCTGCAATGGCTGTGCTC	1020
QY	1034	AGCTGTGTCTGAGATGTTGGGTTAAATCCCGCAACGAGCGCAACCTTGTCTATTAT	1093
Db	1021	AGCTGTGTGTGAATATGTGGGTTAAATCCCGCAACGAGCGCAACCTTATCTTGTGTT	1080
QY	1094	GCCATCA-TTTGGTTGGGCACTTTAATGAACCTGCGGTGCAAAACGGAGAAAGGTGGG	1152
Db	1081	GCCACGGGTTCCGCGGGAACTCAAAAGAGACTGCATGTATTAATCTGAAGAAAGGTGGG	1140
QY	1153	GATGACGTCAAGTCTCAATGCGCCCTTAATGGGATGGGCTTCAACAGTAAATCAATGCGCG	1212
Db	1141	GATGACGTCAAGTCTCAATGCGCCCTTAAGAGTGGGCTTCAACAGTGTCTCAATGGGCGCA	1200
QY	1213	TACAGAGGGTCCCAACCCGCGAGGGGAGCTAATCTCAAAAGCGCGTGTAGTCCGGA	1272
Db	1201	TACAAAGAGAACCGACTTCGCGAGAGCAAGCGGAACTTAAAGTGCCTGTAGTCCGGA	1260
QY	1273	TGCGAGTCTGCAACTGCACTCCGTGAAGTCGGAATCGCTAGTAATCGGGATCCAGACTGT	1332
Db	1261	TTGGAATCTGCAACTGCACTCCATGAAGTCGGAATCGCTAGTAATCGGGATCCAGATGC	1320
QY	1333	CGGGGTGAATACGTTCCCGGCTCTTGTACACACCGCCGTCACACATGGAGTGGGTTT	1392
Db	1321	CAGCGTAATACGTTCCCGGCTCTTGTACACACCGCCGTCACACATGGAGTGGGTTG	1380
QY	1393	CACCAAGACAGGTACTTAACCGTAAAGAGGGCGCTTGCCAGGTTGAGATTCAATGCTG	1452

Db	1361	CAAAAGAGTAGTAGTCTTAACCTTCGGAGGCGCTTACCACTTTGTGATTCATGACTG	1440
Qy	1453	GGGTG 1457	
Db	1441	GGGTG 1445	

RESULT 34  
US-10-831-286A-48689/c

Sequence 48689, Application US/10831286A  
Publication No. US20060046246A1  
GENERAL INFORMATION:  
APPLICANT: ZENG, QIAN DONG  
APPLICANT: CHATELIER, SONIA  
APPLICANT: MOIR, DONALD T.  
APPLICANT: LACROIX, BRUNA  
APPLICANT: CHILDRESS, DARELL  
TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
FILE REFERENCE: 032796-174.001  
CURRENT APPLICATION NUMBER: US/10/831, 286A  
CURRENT FILING DATE: 2004-04-26  
PRIOR APPLICATION NUMBER: 60/464, 955  
PRIOR FILING DATE: 2003-04-24  
NUMBER OF SEQ ID NOS: 48788  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 48689  
LENGTH: 1525  
TYPE: DNA  
ORGANISM: Enterobacter hormaechei

US-10-831-286A-48689

Query Match 67.1%; Score 977.8; DB 11; Length 1525;  
Best Local Similarity 82.4%; Pred. No. 1,7e-293;  
Matches 1194; Conservative 0; Mismatches 247; Indels 8; Gaps 6;

Qy	17	ATGCTTTAACAATGCAAGTCGAAACGGCAGACG--GATGCTTCATCT--GATGCGAGT	72
Db	1510	AGGCTTAACAATGCAAGTCGAAACGGCAGACGCTTGTCTGCTTGGCTGACGAGT	1451
Qy	73	GGCGGACGGGTGAGTAAATGATCGAAGCCTATCAGAAAGGCGGGTAAACGATCGAAG	132
Db	1450	GGCGGACGGGTGAGTAAATGATCGAAGCCTATCAGAAAGGCGGGTAAACGATCGAAG	1391
Qy	133	ATGTGCTAAATACCGCATATATCTTAAAGAGAAAGCAGGGGATCGAAAGACTTGGCTT	192
Db	1390	GGTAGCTAATACCGCATATATCTTAAAGAGAAAGCAGGGGATCGAAAGACTTGGCTT	1331
Qy	193	TTGGAGCGCCGATATCTGATTTAGTGAATGATGGGTTAAAGGCTTACCAAGCGACGAT	252
Db	1330	TCGAGTGTGCCGAGATGGAATTAAGTAAAGTGAAGGCTTACCGCTCACTAGGCGACGAT	1271
Qy	253	CAGTAGTGGTCTGAGAGGACGACCACTGGGACTGAGACAGCGCCAGACTTCT	312
Db	1270	CCCTAGCTGCTCTGAGAGGATGACAGCCACACTGGAATCGAGACGCTCAGACTTCT	1211
Qy	313	ACGGAGGACGAGTGGGAAATTTTGACATATGGGCCAAGCTGATTCAGCAATGCGC	372
Db	1210	ACGGAGGACGAGTGGGAAATTTTGACATATGGGCCAAGCTGATTCAGCAATGCGC	1151
Qy	373	GTGAGTGAAGAAGGCTTGGGTTGTAAAGCTCTTTCAGTCSAGAAAGAAAGTTACGCT	432
Db	1150	GTGATGAAGAAGGCTTGGGTTGTAAAGCTCTTTCAGTCSAGAAAGAAAGTTACGCT	1091
Qy	433	AAATAATTCGACCCATGACGATATGACAGAAAGACCCGCTAACTACGTGACGCA	492
Db	1090	TAAATACTCAGCAATGACGATATGACAGAAAGACCCGCTAACTACGTGACGCA	1031
Qy	493	GCCGCGGTAATAGTAGGGTGCAGAGCTTATCGAATTTCTGGGCGTAAAGGGTGC	552
Db	1030	GCCGCGGTAATAGTAGGGTGCAGAGCTTATCGAATTTCTGGGCGTAAAGGGTGC	971
Qy	553	GGGCGGCTTAAAGTCAATGTGAAATCCCGGGCTTAACTCGGGAATTTGCTTGAAC	612

Db 970 GGGGCTCTGCAAGTCGATGTGAATCCCGGGCTCACTGGGAATGCATTCGAAAC 911  
Qy 613 TACAAGTAGAGTGTGGCAGAGAGGTGGAATTCATGTGATGACATGAATGGGTAG 672  
Db 910 TGGCAGGCTGAGCTCTGTGTAGAGGGGGGTGAATTCAGAGGTGTAGCGGTGAATGGGTAG 851  
Qy 673 AGATATGGAAGAATCATCGATGGGGAAGGCGCTCCGGGTAACTGACGCTCATGCA 732  
Db 850 AGATCTGAGGAATACCGGTGGGAGAGCGCCCGCTTGACAAAGACTGACGCTCAGGTG 791  
Qy 733 CGAAAGCTGGGAGAGCAACAGATTAGATACCTGTAGTCCACGCGCTTAAACGATGTC 792  
Db 790 CGAAAGCTGGGAGAGCAACAGATTAGATACCTGTAGTCCACGCGCTTAAACGATGTC 731  
Qy 793 AACTAGTTGTGGGCTTATTAGGCTTGG--TAAAGAACTTACGCGTGAATGTTACCGCC 851  
Db 730 GACTTGGAGGTTGTGCTTGAAGCGCTTGCGCTTCGAGCTAAGCGCTTAAAGTGCAGCGCC 671  
Qy 852 TGGGAGTAGAGTTCGCAAGATTAAACTCAAGAAATTGACGGGGAACCGCAGACGGGT 911  
Db 670 TGGGAGTAGAGTTCGCGCAGGATTAAACTCAAGTAATGACGGGAGCGCCGACAGCGGT 611  
Qy 912 GGAATTATGTGATTAATTGATGCAACGCGAAAACTTACTACCTTGAACATGTAGCG 971  
Db 610 GGAAGCATGTGTTAATTGATGCAACGCGAACGTTACTTACTTGTACATCCAGAG 551  
Qy 972 AATTTCTAGAGATGATTAATGAGTGTG--CTTGGGAAACGCTTACACAGAGTGTGCTGTC 1030  
Db 550 AACTAGCAGAGATGCTTGTGTGCTTGGGAACTGTGAGACAGGTGCTGACATGCGTGTTC 491  
Qy 1031 GTAGCTCGTGTGTGATGATGTTGGGTAGTCCCGCAAGAGCGCAACCTTGTGCTA 1090  
Db 490 GTAGCTCGTGTGTGATGATGTTGGGTAGTCCCGCAAGAGCGCAACCTTATCTT 431  
Qy 1091 ATTGCATCA--TTTGGTTGGGCACTTTAATGAGATGCGCGGTGACAAACCGAGGAAGT 1149  
Db 430 GTTGCAGAGCGTTAAGCGCGGAACTCAAGAGACATGCCAGTATTAATCGAGGAAGGT 371  
Qy 1150 GGGGATGACGTCAAGTCTCATGAGCCCTTATGAGGAGGCTTACACAGTATTAATGAGT 1209  
Db 370 GGGGATGACGTCAAGTCTCATGAGCCCTTACGATGAGGCTTACACAGTCTCAATGAGT 311  
Qy 1210 GCGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAAGAAAGCGGTGTGATGCC 1269  
Db 310 GCAATCAAGAAAGAGGCACTCGGAGAGCAAGCGGACTCATTAAGTGTGATGATCC 251  
Qy 1270 GGAATGGAAGTCTGCAACTGCACTCCGTAAGTGGAAATCGCTAGTATCGCGATGACGA 1329  
Db 250 GGAATGGAAGTCTGCAACTGCACTCCGTAAGTGGAAATCGCTAGTATCGTGAATCAGAA 191  
Qy 1330 TGTGCGGTGATTAAGTCTCCGGGCTTGTGTAACAACCGCGCTTACACCATGGAAGTGG 1389  
Db 190 TGCACAGGTGATTAAGTCTCCGGGCTTGTGTAACAACCGCGCTTACACCATGGAAGTGG 131  
Qy 1390 TTTTCAACAAGAGAGTATGTTAAACCGTAAAGAGAGGCGCTT--GCGACGATGATTCATG 1448  
Db 130 TTGCAAAAAGAGTATGATTAACCTTGGGAGGGGCTTACCACTTTGTATTCATG 71  
Qy 1449 ACTGGGCTG 1457  
Db 70 ACTGGGCTG 62

RESULT 35  
US-11-055-637-76  
; Sequence 76, Application US/11055637  
; Publication No. US20050260619A1  
; GENERAL INFORMATION:  
; APPLICANT: BROUSSEAU, Roland  
; APPLICANT: DUBOIS, Jason  
; APPLICANT: EDGE, Tom  
; APPLICANT: MASSON, Luc

; APPLICANT: TREVORS, Jack T.  
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND  
; TITLE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES  
; FILE REFERENCE: 2139-33US  
; CURRENT APPLICATION NUMBER: US/11/055,637  
; PRIOR FILING DATE: 2005-02-11  
; PRIOR APPLICATION NUMBER: US 60/543,288  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 76  
; LENGTH: 1485  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Probe for DNA array  
US-11-055-637-76

Query Match 58.5%; Score 852; DB 17; Length 1485;  
Best Local Similarity 76.1%; Pred. No. 2,4e-254;  
Matches 1119; Conservative 0; Mismatches 335; Indels 16; Gaps 5;

Qy 4 GAACGCTGGCGCATGCTTTTACACATGCAAGTCGAACGCGACGACGATGCTTGCATCTG 63  
Db 4 GAACGCTGGCGCATGCTTTTACACATGCAAGTCGAACGCGACGATGCTTGCATCT 63  
Qy 64 GTGGGAGTGGGGGACGGGTGATGATGATCGG--AAGTATCCAAAGAGGGGGGTAC 122  
Db 64 GATTTTACGGGGGACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 123  
Qy 123 GCATGAAAGATGTCTAATACCGCATTAATCTTAAGAGAGAAACAGGGGATTCGAA-- 180  
Db 124 TCCGGGAAACCGGGGCTAATACCGCATTAATCTTAAGAGAGAAACAGGGGATTCGAA 183  
Qy 181 -----GACCTTGGGCTTTTGGAGCGCGCATGCTGATTAAGTGAATGTTGGGGTAA 233  
Db 184 ACGGTTTCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243  
Qy 234 GGCCTTACCAAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 293  
Db 244 GGCCTTACCAAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303  
Qy 294 AGACAGCGCCAGACTCTCTACGAGGAGGAGCAGTGGGAAATTTTGAACATGGGCGCAAG 353  
Db 304 AGACAGCGCCAGACTCTCTACGAGGAGGAGCAGTGGGAAATTTTGAACATGGGCGCAAG 363  
Qy 354 CCTGATCCAGCAATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 413  
Db 364 TCTGATGAGCAATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 423  
Qy 414 GAGAAAGAAAGTTACGTTAATTAATCGTGAACCCATGACCGTATGACAGAAAGACACC 473  
Db 424 GAGAAAGAAAGTTACGTTAATTAATCGTGAACCCATGACCGTATGACAGAAAGACACC 483  
Qy 474 GGTAACTACGTGCAAGAGCGCGGTAAATAGTGGGTGCAAGCGTTAATCGAATTAAC 533  
Db 484 GGTAACTACGTGCAAGAGCGCGGTAAATAGTGGGTGCAAGCGTTAATCGAATTAAC 543  
Qy 534 TGGGCTGAAGAGGTGGGAGGCGCTTGTAAAGTCAAGTGAATATCCCGGGCTTAAAC 593  
Db 544 TGGGCTGAAGAGGTGGGAGGCGCTTGTAAAGTCAAGTGAATATCCCGGGCTTAAAC 603  
Qy 594 TGGGAATTCGTTTAAATTAACAAAGCTGATGATGATGATGATGATGATGATGATGAT 653  
Db 604 GTGGAGGCTCATTTGAAACTGGAAGACTTGAATACAGAAAGAGAAAGCGAATTCACGT 663  
Qy 654 GTAGCAGTGAATGCGTAGAGATATGAAAGACATGATGCGAAGGCGAGCTCTGGGT 713  
Db 664 GTAGCAGTGAATGCGTAGAGATATGAAAGACATGATGCGAAGGCGAGCTCTGGGT 723  
Qy 714 TAAACATGACGCTTACGCAAGAACGTTGGGAGGCAACAGGATTAAGTATCCCTGTAGT 773  
Db 724 TGTAACTGACGCTTACGCAAGAACGTTGGGAGGCAACAGGATTAAGTATCCCTGTAGT 783

QY	774	CAAGCCCTTAACGATGTCAACTAGTTGTT--GGGCTTATTAGGCTGGTAACGAGCT	831
Db	784	CCACCCCGTAAACGATGAAGTCTAAAGTGTATGGGAGGTTTCCGCCCTTAAGTGTGACGCT	843
QY	832	AAAGGCTGAAGTTGAACCGCCTGGGAGAGTACGGTCCGACAGATTAAACTCAAGAAATTGA	891
Db	844	AACGATTAAAGCACTCCGCTTGGGAGTACGCGCCGCAAGGCTGAACCTCAAGAAATTGA	903
QY	892	CGGGGACCCGCAACAAGCGGTGGATTATGTGATTAAATTGATSCAACGCGAAAACTTGA	951
Db	904	CGGGGACCCGCAACAAGCGGTGGAGCATGTGTGTTTAATTGCAACCAACGCAAGAACTTGA	963
QY	952	CCTAACCTTGAACATGTAGCGAA--TTTCTAAGATAGTATAGTGTTCGGGAACCTGA	1008
Db	964	CCAGGTCCTTGAACATCCCACTGAACCGGTGTAGAAATACGCCCTTCCCTTCGGGACAACTGG	1023
QY	1009	ACACAGTGTCTCATGCGCTGTGTGTCAGCTGTGTGCGAGATGTTGGGTTAAGTCCGCA	1068
Db	1024	TGACAGGTGTGATGTTGTGTGTGTCAGCTGTGTGTGCGAATGTTGGGTTAAGTCCGCA	1083
QY	1069	ACGAGCGCAACCTTGTGTCAATTAAATTGCCATCATTTGGTTGGGCACTTAAATGACATGCC	1128
Db	1084	ACGAGCGCAACCTTGTATCTTAAGTGTCCAGCATTTCAAGTTGGGCACTTAAATGATGCC	1143
QY	1129	GGTGCACAAACCGGAGAGGTGGGATGACGTCAAGTCTCATGAGCCCTTATGGGTAGGG	1188
Db	1144	GGTGATTAACCGGAGAGAGGTGGGATGACGTCAAAATCATCATGACCCTTATGACTGGG	1203
QY	1189	CTTCACACGTAAATACATATGGCCGCTACAGAGGTTGTCACCCGCGAGGGGGAGCTATC	1248
Db	1204	CTACACACTGTCTACATATGATGATACAGAGGTTGTCACCCGCGAGGGGGAGCCATC	1263
QY	1249	TCAGAAAGGCGGTGATGTCCGATGCGGATGTGCAACTGCAGCTCCGTGAAGTGGGATC	1308
Db	1264	CCATTAATGTCTCCAGTTCGAGTTTGGAGAGCTGCACCTGCCTTCATGAAGTTGATC	1323
QY	1309	GCTAGTAATCGCGGATCAGCATGTCCGCGTGAATACGTTCCCGGCTTGTACACACGC	1368
Db	1324	GCTAGTAATCGTGGATCAGCATGCACAGGTAATACGTTCCCGGCTTGTGACACACGC	1383
QY	1369	CCGTCAACCAATGGAGTGGGTTTCAACCAAGACAGTACTTA-ACCCTAAGAGGGCG	1427
Db	1384	CCGTCAACCAACGAGATTGTGAACCCCAAGTCGTTGGGGTATCATCTACGAGAGCCAG	1443
QY	1428	CTTGCCACGCTGATTCATGACTGGGCTG	1457
Db	1444	CCGCCAAGGTGGACAGATGATTGGGGTG	1473

```

/ RESULT 36
/ US-11-055-637-80
/ Sequence 80, Application US/11055637
/ Publication No. US20050260619A1
/ GENERAL INFORMATION:
/ APPLICANT: BROUSSEAU, Roland
/ APPLICANT: DUBOIS, Jason
/ APPLICANT: EDGE, Tom
/ APPLICANT: MASSON, Luc
/ APPLICANT: TREVORS, Jack T.
/ TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
/ TITLE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES
/ FILE REFERENCE: 2139-33US
/ CURRENT APPLICATION NUMBER: US/11/055,637
/ CURRENT FILING DATE: 2005-02-11
/ PRIOR APPLICATION NUMBER: US 60/543,288
/ PRIOR FILING DATE: 2004-02-11
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 80
/ LENGTH: 1505
/ TYPE: DNA
/ ORGANISM: Artificial Sequence

```

Query Match	58.2%	Score 847.4	DB 17	Length 1505
Best Local Similarity	75.9%	Pred. No. 6.6e-253		
Matches 1115	Conservative	0	Mismatches 339	Indels 16
				Gaps 5
US-11-055-637-80				
FEATURE: INFORMATION: Probe for DNA array				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (1) ... (1505)				
OTHER INFORMATION: n = A,T,C or G				
4	GAACGCTGCGCGCATGCTTTACACATGCGAATCGAATCGGACGACGATGCTTGCATCTG	63		
4	GAACTGTCGCGCGCTGCTTATACATGGAATTCGAGCGGACCAAGAGGCTGCTCCCA	63		
64	GTGGCGATGCGCGGACGCGGTGATGATCATCGG- AACGTATCCAGAAAGGGGGGTAA	122		
64	GAGGTATGCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	123		
123	GCATCGAAGATGCTTAAATCCGAT-----AATCTTAAAGGAGAAAGGAGGGG	173		
124	ATCGAAGAAATGCTGCTTAAATCCGATGATGATGATGATGATGATGATGATGATGATGAT	183		
174	ATCGAAGAAATGCTGCTTAAATCCGATGATGATGATGATGATGATGATGATGATGATGAT	233		
184	ATGCTGCTGCTGCTTAAATCCGATGATGATGATGATGATGATGATGATGATGATGATGAT	243		
234	GAGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	293		
244	GAGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	303		
294	AGACACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	353		
304	AGACACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	363		
354	CCTGATCCAGAAATGCGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	413		
364	TCTGACGAGCAACCGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	423		
414	GAGAAGAAAGGTTACGCT- AAATTAATGCTGACCCATGACGCTATCGACAGAAAGACAC	472		
424	GGAAGAAACAACTGCGCTTAAATGCTGACCCATGACGCTATCGACAGAAAGACAC	483		
473	CGGCTAACTACGCTGCGACGACCGCGGTATACGTAAGGTGCAAGCCTTAATCGGAATTA	532		
484	CGGCTAACTACGCTGCGACGACCGCGGTATACGTAAGGTGCAAGCCTTAATCGGAATTA	543		
533	CTGGGCGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	592		
544	TTGGGCGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	603		
593	CTGGGAAATTCGTTGAAATCTCAAAAGCTGAGTGTGCGAGAGGAGGAGGAGGAGGAGGAGG	652		
604	CCCGAGGGGTCAATTGAAATCTGGAGACTTGAATGACAGAAAGGAGGAGGAGGAGGAGGAGG	663		
664	TCTACCGGTGAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	723		
713	TTTAACACTGACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	772		
724	CTGTAACTGACGCTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	783		
773	TCCAGGCGCTTAAGATGCACTGATGTTGTTGGGCTTATGAG--CTTGATTAAGCAAGC	830		
784	TCCAGGCGCTTAAGATGCACTGATGTTGTTGGGCTTATGAG--CTTGATTAAGCAAGC	843		
831	TAAACCGTGAAGTTGACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	890		
844	TAAACCTTAAGCACTCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	903		
891	ACGGGAGCCGCAAGAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT	950		

Db 904 ACCGGGGCCCCGCAACAGCATGGAGCATGTGTTTATTGAAAGCAACGGGAACCTT 963  
Qy 951 ACCTACCTTTGACATGATGAGGAATTTTCTGAGATAGA---TTAGTCTTCGGGAACCT 1007  
Db 964 ACCAGGCTTTGACATCTTTTGACACCTCTAGAGATAGAGCTTTCCCTTCGGGGGACAA 1023  
Qy 1008 AACACAGTGTCTGATGAGTGTGTGATGCTGTGTGTGTGATGTTGGGTTAGTCCCG 1067  
Db 1024 GTGACAGGTGTGATGAGTGTGTGTGATGCTGTGTGTGTGATGTTGGGTTAGTCCCG 1083  
Qy 1068 AACGAGCGCAACCTTTGTCAATTTGATTCATTTGATTTGGGCACTTTATGAGACTGC 1127  
Db 1084 AACGAGCGCAACCTTTGATCTTATGTTCCAGCATTTTATGTTGGGCACTTTAGGTGACTGC 1143  
Qy 1128 CGGTGACAAACCGGAGAGGTGGGATGACGTCAAGTCTTCATGCGCTTTATGGGTAG 1187  
Db 1144 CGGTGACAAACCGGAGAGGTGGGATGACGTCAATTCATGCGCTTTATGACCTGG 1203  
Qy 1188 GCTTCAACGTAATATCAATGGCCGCTACAGAGGTTGCCAACCCGAGAGGGAGCTAAT 1247  
Db 1204 GCTACACACGTGTACATGATGTATCAAAAGGCGACGAAACCGGAGGTGACCAAT 1263  
Qy 1248 CTCAGAAAGCGGTCTGATGCTCGGATCGAGTCTGCAACTCGACTCGTGAAGTCGGAAT 1307  
Db 1264 CCATTAAGCCATTTCTGATTCGATTTGAGCTTGCACACTCGCTTACATGAAACCGGAAT 1323  
Qy 1308 CGGTAGTAATCGCGATGACATGTGCGGTGAATACGTTCCGGGTCTTTTACACACCG 1367  
Db 1324 TGCTAGTAATCGCGATGACATGTGCGGTGAATACGTTCCGGGTCTTTTACACACCG 1383  
Qy 1368 CCCGTACACCATGGAGTGGGTTTCAACGAAAGCGGTAGTCTAACCCGTAAGAGGGCG 1427  
Db 1384 CCGGTACACCAAGAGGTTGTGAACACCCGAAAGTCGTGGGTAACTTTTGAAGCCAG 1443  
Qy 1428 CTGCGACGGTGAGATTGATGACTGGGGTG 1457  
Db 1444 CCGCTTAAGGTGGGACGATGATTGGGGTG 1473

## RESULT 37

US-11-055-637-79  
; Sequence 79, Application US/11055637  
; Publication No. US20050260619A1  
; GENERAL INFORMATION:  
; APPLICANT: BROUSSEAU, Roland  
; APPLICANT: DUBOIS, Jason  
; APPLICANT: EDGE, Tom  
; APPLICANT: MASSON, Luc  
; APPLICANT: TREVORS, Jack T.  
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND  
; FILE REFERENCE: 2139-3105  
; CURRENT APPLICATION NUMBER: US/11/055,637  
; PRIOR FILING DATE: 2005-02-11  
; PRIOR APPLICATION NUMBER: US 60/543,288  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 79  
; LENGTH: 1513  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Probe for DNA array  
US-11-055-637-79

Query Match 58.1%; Score 847.2; DB 17; Length 1513;  
Best Local Similarity 77.2%; Pred. No. 7.6e-253;  
Matches 1155; Conservative 0; Mismatches 318; Indels 17; Gaps 8;

Qy 5 AACGTGGCGGCAATGCTTACATGCAATGCAAGCGACGACGATGCTTGCATCTGG 64  
|||||

Db 1 AACGTGGCGGCGCTTAATATCATGCAAGTCGACGAAATGGAATTAAGACTTGTCTTTA 60  
Qy 65 TG--GGAGTGGCGGACGGGTGATGATGATCATCGG-AACGTATCCGAAAGAGGGGGTTAA 121  
Db 61 TGAAGTTACGGCGGAGCGGTGATGACAGTGGGTAACTTACCATTAAGATCTGGGATTA 120  
Qy 122 CGCATGAAAGATGTGTTAATACCGCAT-ATATCTTAAGAGGAAAGCAGGGGATCGAAA 180  
Db 121 CTCGGGAAACCGGGGCTTAATACCGGATTAATTTTGAACATGATAGTTGAAATTGAAA 180  
Qy 181 GAC-----CTTGGCTTTTGGAGCGGCGCATGTCTGATTTAGTACTGTTGGTGGGTTA 232  
Db 181 GCGGCTTGGCTTCACTTATGATGACGCGCGCTTCGATTAAGCTAGTTGGTGAAGTTA 240  
Qy 233 AAGCTTACCAAGCCGACCATTCAGTGTGTTGTTGAGAGGACGACCAACGACCTGGGACT 292  
Db 241 CCGCTTACCAAGCCGACCATTCAGTGTGTTGTTGAGAGGATATCGGCACACTGGGACT 300  
Qy 293 GAGACACGCGCCGAGACTCTTACGAGAGCAGAGTGGGGAATTTTGAACATGGCGCA 352  
Db 301 GAGACACGCGCCGAGACTCTTACGAGAGCAGAGTGGGGAATTTTGAACATGGAGAA 360  
Qy 353 GCTGATCCAGCAATGCGCGTGAAGTGAAGAAAGCCCTTCGGGTGTTAAAGCTCTTCACT 412  
Db 361 GTCTGACGAGCAACGCGCGTGAAGTGAAGAAAGCCCTTCGGGTGTTAAAGCTCTTCACT 420  
Qy 413 CGAAGAAAGAAAGTTACGTTA--AATATCTGACCCATGACGTTTTCGACAGAAAGCA 471  
Db 421 AAGGAAAGAAAGTTACGTTAATTAAGTGTGACCTTGAACGTTACTTAACAGAAAGCC 480  
Qy 472 CCGCTTAATCACTGCTCCAGACACCGCGTAAATATAGAGGTGCAAGCGTTAATTCGAAT 531  
Db 481 AAGCTTAATCACTGCTCCAGACACCGCGTAAATATAGAGGTGCAAGCGTTAATTCGAAT 540  
Qy 532 ACTGGCGGTAAAGGTGCGCAGCGCGCTTGTAACTGATGATGAAATCCCGGCTTAA 591  
Db 541 ATTGGCGGTAAAGGTGCGCAGCGCGCTTGTAACTGATGATGAAATCCCGGCTTAA 600  
Qy 592 CCGGGAATTTGGCTTTGAAACTTAAGTATGATGATGATGATGATGATGATGATGAT 651  
Db 601 CCGGGAATTTGGCTTTGAAACTTAAGTATGATGATGATGATGATGATGATGATGAT 660  
Qy 652 GTGTAGCAGTGAATGCGTGAAGATATGAGAAATCATGATGAGCAAGCAGCTCTCTGG 711  
Db 661 GTGTAGCAGTGAATGCGTGAAGATATGAGAAATCATGATGAGCAAGCAGCTCTCTGG 720  
Qy 712 GTTAACTATGACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGATACCTGTGTA 771  
Db 721 TCTGTAACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGATACCTGTGTA 780  
Qy 772 GTTCAAGCCCTTAAGCATGATCAATGATTTGGGCTTTATTAAGC--TTGGTAAAGAG 829  
Db 781 GTTCAAGCCCTTAAGCATGATGATGATTTGGGCTTTATTAAGC--TTGGTAAAGAG 840  
Qy 830 CTAAAGCGTGAAGTTGACCGGCTGGGAGTAAAGTGCAGAAATTAAGCTCAAGAGAAAT 889  
Db 841 TTAAGCATTAAGCACTCGGCTGGGAGTAAAGTGCAGAAATTAAGCTCAAGAGAAAT 900  
Qy 890 GAGGGGAAACCGGCAAGAGCGGTGATTAATGATGATTAATTCGATCAACGCAAAACCT 949  
Db 901 GAGGGGAAACCGGCAAGAGCGGTGATTAATGATGATTAATTCGATCAACGCAAAACCT 960  
Qy 950 TACCTAATCTTAATGATGAGCAATTTTGTAGAGATAGA-TTAAAGTCTTCGGGAACGCTA 1008  
Db 961 TACCAAGTCTTAATGATGAGCAATTTTGTAGAGATAGA-TTAAAGTCTTCGGGAACGCTA 1020  
Qy 1009 AACAGGTGTGATGAGTGTGTGATGCTGTGTGTGATGCTGTGTGTGATGCTGTGTGAT 1068  
Db 1021 TGAAGTGTGATGAGTGTGTGATGCTGTGTGTGATGCTGTGTGTGATGCTGTGTGAT 1080  
Qy 1069 AAGAGCGCAACCTTGTCAATTAATGATGATTAATTTGGTGGGCACTTTAATGAGACTGCC 1128  
Db 1081 AAGAGCGCAACCTTGTCAATTAATGATGATTAATTTGGTGGGCACTTTAATGAGACTGCC 1140

QY 1129 GTGACAAACCGAGGAGGTGGGATGACGTCAAGTCTTCATGCGCCCTTATGCGTAAAGG 1188  
DB 1141 GGTGACAAACCGAGGAGGTGGGATGACGTCAAAATCATCATGCCCCCTTATGACCTGGG 1200  
QY 1189 CTTACACACCTAATATCAATGCGGGTATAGAGGGGTGCGCAACCGCGAGGGGGAGCTAATC 1248  
DB 1201 CTACACACGTGTCAATATGACCGGTACAAAGAGCTGCAAGACCGCGAGGTGAGCTAATC 1260  
QY 1249 TCAGAAAGCGCGTGTAGTCCGGATCCGAGTCTGCACTCGACTCCGTGGAAGTCGAATC 1308  
DB 1261 TCAATAAACCGCTTCTCAGTTCCGATTTAGGCTGCAACTCGCTCAKATGAAGCTGGAATC 1320  
QY 1309 GCTAGTAATCGCGGATCAGCATGTCCGGGTGAATACGTTCCGGGCTTTGTACACACCGC 1368  
DB 1321 GCTAGTAATCGCGGATCAGCATGCGCGGTGAATACGTTCCGGGCTTTGTACACACCGC 1380  
QY 1369 CCGTCACACCAATGGGATGGGTTTACACAGAGAGCTAATCC-GTAAAGAGGGCG 1427  
DB 1381 CCGTCACACCAAGAGGTTTGTACACCCGAAAGTCGGTGGGTAACTTTATGAGCGCAG 1440  
QY 1428 CTTGCCACGGTGAATTCATGATCGGGGTG 1457  
DB 1441 CCGCTTAAGGTGGAGAGATGATGGGGTG 1470

RESULT 38  
US-11-273-617-6  
Sequence 6, Application US/11273617  
Publication No. US20060067924A1  
GENERAL INFORMATION:  
APPLICANT: Lee, Margie D.  
APPLICANT: Hofacre, Charles  
APPLICANT: Harton, Barry  
TITLE OF INVENTION: Probiotic Bacteria and Methods  
FILE REFERENCE: 18-03  
CURRENT APPLICATION NUMBER: US/11/273, 617  
CURRENT FILING DATE: 2005-11-14  
PRIORITY APPLICATION NUMBER: US 60/470,807  
PRIORITY FILING DATE: 2003-05-14  
PRIORITY APPLICATION NUMBER: NO us2004/015378  
PRIORITY FILING DATE: 2005-04-14  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 6  
LENGTH: 1509  
TYPE: DNA  
ORGANISM: Enterococcus cecorum  
US-11-273-617-6

Query Match 57.8%; Score 842.8; DB 18; Length 1509;  
Best Local Similarity 77.9%; Pred. No. 1.8e-251;  
Matches 1091; Conservative 0; Mismatches 298; Indels 12; Gaps 6;

QY 69 GAGTGGCGAGCGGGTGAATATGATCGG-AACGTATCCAGAGAGGGGGGTAAACGATC 127  
DB 81 GAGTGGCGAGCGGGTGAATATGATCGGATTAACGTCGATCAAGGGGGGTAAACATTTG 140  
QY 128 GAAAGATGTCTAATACCGCAT-ATACTCTAAGAGAGAAAGCAGGGGATCGAAAGACTT 186  
DB 141 GAAACAGGTCTAATACCGCATTAATTCATTTACCGCATGTGATGATTAAGAGCGCT 200  
QY 187 G-----CGCTTTTGGAGCGGCGCATGTCTGATTAGTAGTATGGGGGTAAAGGCTTAC 240  
DB 201 TTTCGCTACTGATGATGATGACCGCGGTGCTTTAGTATGTTGGGTAAACGCTTAC 260  
QY 241 CAAGGCGAGATCACTAGTATGCTGTGAGAGAGCAACGACCACTGAGGACTGAGACAG 300  
DB 261 CAAGGCTGAGTATGATGACCGACCTGAGAGGGGTGATCGGCACTGAGGACTGAGACAG 320  
QY 301 GCCCAGACTCTTAACGGAGGAGCAGCTGGGAAATTTGACAAATGGGCGAAGCTGATC 360  
DB 321 GCCCAGACTCTTAACGGAGGAGCAGCTGAGGAAATCTTCGGCAATGAGCAAGCTGAC 380

QY 361 CAGCAATGCCGCGTGAATAGAGAGGCTTCGGGTGTAAAGCTTTCACTGAGAGAGA 420  
DB 381 GAGCAACGCGCGGTGAATAGAGAGGTTTCCGATGTAAATCTCTGTTGTAAGAGAGA 440  
QY 421 A-AAGGTTACGGTAAATTAATCGTGAACCATGACGATATGACAGAGAGAGCAGGCTTAA 479  
DB 441 ACAAGATGAGAGTGAAGATTATCTGACGATATCTAACAGAAAGCCAGGCTTAA 500  
QY 480 CTACGTCCAGACCGCGGTAAATACGTAGGGTGCAGAGGCTTAAATGGAATTAATGCGGC 539  
DB 501 CTACGTCCAGACCGCGGTAAATACGTAGGGTGCAGAGGCTTAAATGGAATTAATGCGGC 560  
QY 540 TAAAGGTTGCGAGCGCGCTTGTAACTGAGATGTAATATCCCGGGCTTAACTGAGAA 599  
DB 561 TAAAGGAGGCGAGCGCGCTTGTAAATGTAATGTAATGTAATGTAATGTAATGTAATG 620  
QY 600 TTGCGTTTGAATCAACAAAGCTAGAGTGGCGAGAGGAGTGAATTCATGTGTAGCA 659  
DB 621 GGTCAATGGAAATCTGGGAGCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 680  
QY 660 GTGAATGCGTGAAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719  
DB 681 GTGAATGCGTGAAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 740  
QY 720 TGAAGCTATGACAGAAAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779  
DB 741 TGAAGCTATGACAGAAAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 800  
QY 780 CTTAAAGAGTGTCAATAGTGTGTTGGGCTTATTAAGGCTTG--GTAAAGAGCTTAACGCG 837  
DB 801 GGTAAAGAGTGTCAATAGTGTGTTGGGCTTATTAAGGCTTG--GTAAAGAGCTTAACGCG 860  
QY 838 TGAAGTGAACCGCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897  
DB 861 TTAAGACATCTCGCTGGGAG 920  
QY 898 CCGGACAGAGCGGTGATTAATGAGATTAATTTGAGTGAACCGGAAACCTTAACCTAAC 957  
DB 921 CCGGACAGAGCGGTGATTAATGAGATTAATTTGAGTGAACCGGAAACCTTAACCTAAC 980  
QY 958 CTTGACATGTAAGAGAAATTTCTAGAGATA-GATTAGTCTTCGGGAAACGCTTAACACAGT 1016  
DB 981 CTTGACATGTAAGAGAAATTTCTAGAGATA-GATTAGTCTTCGGGAAACGCTTAACACAGT 1040  
QY 1017 GCTGATGCTGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1076  
DB 1041 GCTGATGCTGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100  
QY 1077 AACCTTGTCAATTAATGCAATCAATTTGTTGGGCACTTAAATGAGACTGCGGTGACAA 1136  
DB 1101 AACCTTGTCAATTAATGCAATCAATTTGTTGGGCACTTAAATGAGACTGCGGTGACAA 1160  
QY 1137 ACCGAGAGAGAGGAGATGACAGTCAAGTCTCAATGCGCTTAATGAGGCTTACGCTTCA 1196  
DB 1161 TGCAGAGAGAGGAGATGACAGTCAATGCTCAATGCGCTTAAATGAGGCTTACGCTTCA 1220  
QY 1197 GTAATCAATGAGCGCTGACAGAGGTTGCCAACCCGAGAGGAGAGGATTAATTCAGAAAG 1256  
DB 1221 GTGCTCAATGAGAGAGTACAAAGAGTGCAGAAAGCGGAGAGCTTAAGCAATCTCTTAAAG 1280  
QY 1257 CGCGTGTAGTCCGGATCGAGATGCTGCACTCGGTAAGTCCGTAATGCTGTAATGA 1316  
DB 1281 CTTCTTCAAGTTCCGATTTGATGCTGCAATCTCGCTTCAATGAAACCGGAAATGCTTAAAG 1340  
QY 1317 TTGCGGATTCAGAGATGCGCGGTGAATGATTTCCGGGCTTTGTACACACCGCGCTTACA 1376  
DB 1341 TTGCGGATTCAGAGATGCGCGGTGAATGATTTCCGGGCTTTGTACACACCGCGCTTACA 1400  
QY 1377 CCATGAGAGTGGGTTTCAACAGAGAGAGTATGTTAAACGTTAAAGAGGCGCTTTCACAG 1436  
DB 1401 CCAAGAGATTTTGAACACCAACCGAGGCGGTAAACCGGAAAGAGCAGCGCTTAAAG 1460



QY 1437 GTGAGATTCATGACTGGGTC 1457  
 DB 1461 GTGGAGTAAGTATGGGGTG 1481

RESULT 39

US-11-055-637-66  
 ; Sequence 66, Application US/11055637  
 ; Publication No. US20050260619A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BROUSSEAU, Roland  
 ; APPLICANT: DUBOIS, Jason  
 ; APPLICANT: EDGE, Tom  
 ; APPLICANT: MASSON, Luc  
 ; APPLICANT: TREVORS, Jack T.  
 ; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND  
 ; FILE REFERENCE: 2139-33US  
 ; CURRENT APPLICATION NUMBER: US/11/055,637  
 ; PRIOR FILING DATE: 2005-02-11  
 ; PRIOR APPLICATION NUMBER: US 60/543,288  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 66  
 ; LENGTH: 1545  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Probe for DNA array  
 US-11-055-637-66

Query Match 57.7%; Score 840.2; DB 17; Length 1545;  
 Best Local Similarity 75.6%; Pred. No. 1.2e-250;  
 Matches 1111; Conservative 0; Mismatches 343; Indels 15; Gaps 5;

QY 4 GAACGCTGGCGGATGCTTTACATGCAAGTGAACCGGACGAGACCGGATGCTGCAATCTG 63  
 DB 24 GAACGCTGGCGGATGCTTTACATGCAAGTGAACCGGACGAGACCGGATGCTGCTCT 83  
 QY 64 GTGCGAGTGGCGGACGAGTGAATGATCCG-ACGTAATCAGAGAGGCGGTAAC 122  
 DB 84 TAGGTGAGCGCGGACGAGTGAATGATCCG-ACGTAATCAGAGAGGCGGTAAC 143  
 QY 123 GCATCGAAGATGCTTAATCCGATATCTTAAGAGAG-AGCAAGG 172  
 DB 144 TCCGGGAAACCGGAGCTTAATCCGATATCTTAAGAGAG-AGCAAGG 203  
 QY 173 GATCGAAGACCTTGGGCTTTGAGAGCGGCGATGCTGATTAAGTGAATGCTGAGGTA 232  
 DB 204 GTGGCTTTTACCTACCTTAACATGAGACCGGCGGCTTAAGTGAATGCTGAGGTA 263  
 QY 233 AGGCTTAACGAAGCGAGATCAATGATGCTGAGAGAGCAACGACCACTGGGACT 292  
 DB 264 CGGCTACCAAGCGAGATGATGATGAGCGGAGGATGAGGCTGAGGAGCT 323  
 QY 293 GAGACAGCGCGGAGACTCTTAACGAGAGAGCAAGTGGGAAATTTTGAACAATGGCGCA 352  
 DB 324 GAGACAGCGCGGAGACTCTTAACGAGAGAGCAAGTGGGAAATTTTGAACAATGGCGCA 383  
 QY 353 GCGTATCCAGCAATGCGGAGTGAAGAGGCGCTTGGGTTGTAAGCTCTTCACT 412  
 DB 384 GTCTGACGAGCAACGCGGAGTGAAGAGGCTTGGGTTGTAAGCTCTTCACT 443  
 QY 413 CGAAGAGAAAGGTTACGTA-ATAATCGTGAACCATGACGATGACAGAGAAACA 471  
 DB 444 AGGAGAGAAAGGTTACGTTGAAATGAGGCGGTAACCTTGAACGTAACAGAAAGCC 503  
 QY 472 CCGGCTAACTACGTCGAGAGCGCGGTAATAGTAGGTCGAAGCGTTAATCGAAT 531  
 DB 504 ACGGCTAACTACGTCGAGAGCGCGGTAATAGTAGGTCGAAGCGTTAATCGAAT 563  
 QY 532 ACTGGGCTAAAGGTCGAGAGCGGCTTGTAGTGAATGTAATCCCGGCTTAA 591

DB 564 ATTGGCGTAAAGCGGCGGACGCGGTTTCTTAAGTGTGATGAAAGCCCCGCTCAA 623  
 QY 592 CCTGGGAATTCGCTTTGAAACTCAAGCTAAGTGTGCAAGAGAGTGAATTCAT 651  
 DB 624 CCGGGAGGTCATTTGAAACTGCGGAACTTAAGTGAAGAGAGAGTGAATTCAT 683  
 QY 652 GTGTACAGTGAATAGCGTGAAGATATGGAAGAAATCGATGGCGAAGCAGCTCTG 711  
 DB 684 GTGTACGCTGAATAGCGTGAAGATATGGAAGAAATCGATGGCGAAGCAGCTCTG 743  
 QY 712 GTTAACTGACGCTTCAATGCAAGAAAGCTGGGAGCAACAGATTAATACCTGTA 771  
 DB 744 TCTGTAACTGACGCTTCAATGCAAGAAAGCTGGGAGCAACAGATTAATACCTGTA 803  
 QY 772 GTTCAAGCTTAAAGATGCAATTAAGTGTGGGCTTAATTAAGC--TTGTAACGAAG 829  
 DB 804 GTTCAAGCTTAAAGATGCAATTAAGTGTGGGCTTAATTAAGC--TTGTAACGAAG 863  
 QY 830 CTAAAGCGTGAAGTTGACCGCTGGGAGTAAGTGTGCAAGATTAATACTCAAGGAAT 889  
 DB 864 CAAAGCATTAAGCACTCCGCTGGGAGTACGTCGCAAGACTGAATCAAGGAAT 923  
 QY 890 GACGGGACCCGCAAGAGGCTGATTAATGATTAATGATGATGCAAGCAAGCAACT 949  
 DB 924 GACGGGACCCGCAAGAGGCTGATTAATGATTAATGATGATGCAAGCAAGCAACT 983  
 QY 950 TACCTACCTTGAATGATGCAAGATTTTGAAGTA-GATTAGTCTTGGGAAAGCTTA 1008  
 DB 984 TACAGGCTTGAATGATGCAAGATTTTGAAGTA-GATTAGTCTTGGGAAAGCTTA 1043  
 QY 1009 ACACAGGCTGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1068  
 DB 1044 TGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1103  
 QY 1069 ACAGGCGCAACCTTGTCAATTAATGCAATTAATGATGATGATGATGATGATGATG 1128  
 DB 1104 ACAGGCGCAACCTTGTCAATTAATGCAATTAATGATGATGATGATGATGATGATG 1163  
 QY 1129 GGTGACAAACCGGAGAGGAGGAGTGAAGTCAAGTCTCATGAGCCCTTAATGAGG 1188  
 DB 1164 GGTGACAAACCGGAGAGGAGGAGTGAAGTCAAGTCTCATGAGCCCTTAATGAGG 1223  
 QY 1189 CTTCACAGTAAATCAATGAGGAGTGAAGTCAAGTCTCATGAGCCCTTAATGAGG 1248  
 DB 1224 CTTCACAGTAAATCAATGAGGAGTGAAGTCAAGTCTCATGAGCCCTTAATGAGG 1283  
 QY 1249 TCAGAAAGCGTGTGATGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1308  
 DB 1284 CCACAAATCTGATTCAGTTCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1343  
 QY 1309 GCTAGTAATCGGAGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1368  
 DB 1344 GCTAGTAATCGGAGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1403  
 QY 1369 CGGTCAACATGAGGAGGAGTTCACAGAAAGCAGTGTAAACCTGAAGAGGAGGCG 1428  
 DB 1404 CGGTCAACATGAGGAGGAGTTCACAGAAAGCAGTGTAAACCTGAAGAGGAGGCG 1463  
 QY 1429 TTGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1457  
 DB 1464 CGCGAAGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1492

RESULT 40

US-11-273-617-8  
 ; Sequence 8, Application US/11273617  
 ; Publication No. US20060067924A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lee, Margie D.  
 ; APPLICANT: Hofacre, Charles  
 ; APPLICANT: Harmon, Barry  
 ; TITLE OF INVENTION: Probiotic Bacteria and Methods



```
FILE REFERENCE: 18-03
CURRENT APPLICATION NUMBER: US/11/273,617
CURRENT FILING DATE: 2005-11-14
PRIOR APPLICATION NUMBER: US 60/470,807
PRIOR FILING DATE: 2003-05-14
PRIOR APPLICATION NUMBER: WO us2004/015378
PRIOR FILING DATE: 2005-04-14
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.2
SEQ ID NO 8
LENGTH: 1587
TYPE: DNA
ORGANISM: Lactobacillus salivarius
US-11-273-617-8

Query Match      57.7%; Score 840.2; DB 18; Length 1587;
Best Local Similarity 76.8%; Pred. No. 1.2e-250;
Matches 1076; Conservative 0; Mismatches 313; Indels 12; Gaps 4;

QY      69 GAGTGGCGAGCGGTGAGTAATGATCGG-AAAGTATCCAGAGAGGGGGTAAAGCATC 127
      |||
DB      121 GAGTGGCGAGCGGTGAGTAATGATCGGAACTGCTCTTAAAGAGGGGATTAACATTG 180
      |||
QY      128 GAAAGATGCTTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGGATCGAAAGAC---- 183
      |||
DB      181 GAAAGATGCTTAATACCGCATTAATCTTAAGAGATCCATGATCTTAAGTAAAGATGGT 240
      |||
QY      184 ----CTTGGCTTTTGGAGGGGGCGGATGCTGATTAAGTAAAGTGGGGTAAAGGCTTA 239
      |||
DB      241 TCTGCTATGCTTTTGAATGAGACCGCGGCTTAATTAATGATGAGGGGTAACGGCTTA 300
      |||
QY      240 CCAAGGCGAGCATCAGTAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 299
      |||
DB      301 CCAAGTGAATGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
      |||
QY      300 GGGCCAGACTCTTAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
      |||
DB      361 GGTCCAACTCTTAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
      |||
QY      360 CCAAGCAATGCGGTGAGTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
      |||
DB      421 GAGAGCAACCGCGGTGAGTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
      |||
QY      420 AAAAGTTAGCTTAATTAATGATCGGATCCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
      |||
DB      481 AACAGAGAGAGAGTAATGATCGGATCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
      |||
QY      480 CTACGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
      |||
DB      541 CTACGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
      |||
QY      540 TAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
      |||
DB      601 TAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
      |||
QY      600 TTGCGTTTGAAGCTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
      |||
DB      661 GTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
      |||
QY      660 GTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
      |||
DB      721 GTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
      |||
QY      720 TGACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779
      |||
DB      781 TGACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
      |||
QY      780 CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 837
      |||
DB      841 CGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
      |||
QY      838 TGAAGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
      |||
```

```
DB      901 ATAAGATTCGCGCTGGGAGTAGACCGCAAGGTTGAAATCAAGAGATTGACGGGGG 960
      |||
QY      898 CCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
      |||
DB      961 CCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
      |||
QY      958 CTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1016
      |||
DB      1021 CTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
      |||
QY      1017 GCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1076
      |||
DB      1081 GGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
      |||
QY      1077 AACCTTGTGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1136
      |||
DB      1141 AACCTTGTGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
      |||
QY      1137 ACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1196
      |||
DB      1201 ACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
      |||
QY      1197 GTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1256
      |||
DB      1261 GGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
      |||
QY      1257 CCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1316
      |||
DB      1321 CCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
      |||
QY      1317 TGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1376
      |||
DB      1381 TGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
      |||
QY      1377 CCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1436
      |||
DB      1441 CCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
      |||
QY      1437 GTGAGATTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1457
      |||
DB      1501 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1521
      |||

RESULT 41
US-11-055-637-71
; Sequence 71, Application US/11055637
; Publication No. US2005026019A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: DUBOIS, Jason
; APPLICANT: EDGE, Tom
; APPLICANT: MASSON, Luc
; APPLICANT: TREVORS, Jack T.
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
; TITLE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES
; FILE REFERENCE: 2139-3305
; CURRENT APPLICATION NUMBER: US/11/055,637
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,288
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Probe for DNA array
US-11-055-637-71

Query Match      57.6%; Score 838.8; DB 17; Length 1486;
Best Local Similarity 76.0%; Pred. No. 3.1e-250;
Matches 1118; Conservative 0; Mismatches 337; Indels 17; Gaps 6;
```

```
OY 3 TGAAGCTGGCGGCAATGCTTTACACATGCAAGTCCGACG-GCAGCAGCAATGCTTGCAATC 61
Db 3 TGAAGCTGGCGGCGCTGCTTAATACATGCAAGTCCGACGCAACTGATTTGAAACCTTGCTTC 62
OY 62 T- GGTGGGAGAGTGGCGGACCGGGTGAATATGCATGCG- AACGTATTCAGAAAGGCGGGT 119
Db 63 TATGACGTTAGCGGCGGACGGGTGAGTAAACAGTGGGCACTGCTGTGAGACCTGGGAT 122
OY 120 AACGCAATGAAAGATGTGCTAAATACCGCATATCTCTAA-----GAGGAAAGCAG 170
Db 123 AACTTCGGGAAACCGAGGCTTAATACCGGATAGGATCTTCTCTTCATGGGAAATATTTGA 182
OY 171 GGGATCGAAAGACTTGGCGCTTTGGAGCGGCCGATGTCTGATTTAGTACTAGTTGGGGGT 230
Db 183 AAGATGTTTCCGCTATCACTTACAGATGGGCCCGCGGTGCTATGCTAGTTGGTGAAGGT 242
OY 221 AAAGGCTTACCAAGGCGACGATCAGTATGTTGGTCTGAGAGGAGCAGCAGCGCACACTGGGA 290
Db 243 AACGCTCACCAAGGCGACGATGCTGATGCCGACTGAGAGGGGTGATCGGCCACACTGGGA 302
OY 291 CTGAGACACGCGCCAGACTCTTACGAGGAGCAGCAGTGGGAAATTTTGGACAAATGGCGGC 350
Db 303 CTGAGACACGCGCCAGACTCTTACGAGGAGCAGCAGTGGGAAATCTTCGCAATGAGACA 362
OY 351 AAGCTTGATCCAGCAATGCGCGGTGAGTGAAGAAAGGCTTGGGGTTGTAAAGCTCTTTCA 410
Db 363 AAGTGTGACGAGCAACGCGCGGTGATGAAGGGCTTTCGGGTGTAAGAACTCTGTG 422
OY 411 GTGAGAGAAAGGTTACGTTAAATATCGTACCCATGACGATACGACAGAAAGAC 470
Db 423 TTAGGAAAGAAACAGTAAAGTAACTGCTTGTACTTGAACGTAACCTTAACCAAGAAC 482
OY 471 ACCGCTTAACTACGTCGACGACGCGCGTAAATACGTAAGGTGCAAGCGTTAATCGGAT 530
Db 483 CACGCTTAACTACGTCGACGACGCGCGTAAATACGTAAGGTGCAAGCGTTAATCGGAT 542
OY 531 TACTGGGCTAAAGGGTGGCGGCGGCTTGTAACTGACAGTGAATATCCCGGGCTTA 590
Db 543 TATTGGGCTTAAAGGGCGCGGCTTGTAACTGACAGTGAATATCCCGGGCTTA 602
OY 591 ACCTGGGAATGCGTTTGAACCTAACAAAGCTAGAGTGGCGAGGAGGAGTGAATTTCA 650
Db 603 ACCGCTGAGGGTCACTTGGAACTGGGGAACCTTGAATGCAAGAAAGAAAGCGGAATTTCA 662
OY 651 TGTGTAGCAGTGAATGCGTGAAGATATGGAAGAACTATGCGGAAGGACCTCTCTG 710
Db 663 CGTGTAGGGGTGAATGCGTGAAGATATGGAAGAAACACAGTGGCGAAGGCGGCTTTTG 722
OY 711 GGTTAACACTGACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGT 770
Db 723 GTCTTAATCTGACGCTGAGGCGCGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGT 782
OY 771 AGTCAACGCGCTTAAACGATGCACTAGTGTGGGCTTATTAAGGCTTGAAGGCTTGAAGG 828
Db 783 AGTCAACGCGCTTAAACGATGCACTAGTGTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 842
OY 829 GCTTAACGCGTGAAGTGAACCGCTGGGAGTACGCGTGCAGAGATTAAACCTCAAGAGAT 888
Db 843 GCTTAACGCTTAAAGCACTCGCTGGGAGTACGCGTGCAGAGATTAAACCTCAAGAGAT 902
OY 889 TGAAGGAGACCGCGACAGCGGCTGATTAATGTGATTAATTTGATGCAACGGAAGAAC 948
Db 903 TGAAGGAGACCGCGACAGCGGCTGATTAATGTGATTAATTTGATGCAACGGAAGAAC 962
OY 949 TTACCTACCGCTTGAACATGTAAGGAATTTTCTAGAGTAGA---TTAGTCTTGGGAAGC 1005
Db 963 TTACCAAGCTTGTGACATCTCTGACAACTCTAGAGTAGAGGCTTCCCTTGGGGAGACA 1022
OY 1006 CTAAACAGAGTCTGATGGCTGTGCTGAGTCTGATGCTGATGATGTTGGTTAAAGTCCC 1065
Db 1023 GAGTGAACAGAGTGTGATGTTGTGTCAAGTCTGTGTGAGATGTTGGTTAAAGTCCC 1082
```

```
OY 1066 GCAACGAGCGGCAACCTTGTGATTAATTTGCCATCATTTGTTGGGACACTTTAATGAGACT 1125
Db 1083 GCAACGAGCGGCAACCTTGTGATTAATTTGCCATCATTTGTTGGGACACTTTAAGGTGACT 1142
OY 1126 GCGGCTGACAAACCGGAGAAAGTGGGAGTACGTCAAGTCTTCAATGCGCTTATGGSTA 1185
Db 1143 GCGGCTGACAAACCGGAGAAAGTGGGAGTACGTCAAGTCTTCAATGCGCTTATGACT 1202
OY 1186 GGGCTTCAACGATTAATCAATAGCGCGGACAGAGGTTGGCAACCGCGAGGAGGAGCTA 1245
Db 1203 GGGCTTCAACGATTAATCAATAGATGATTAACAAAGGCTGCAAGACCGGAGGCTAAGCA 1262
OY 1246 ATTCAGAAAGCGCGTGTAGTCCGATCGGATCGGATCTGCACTCGCTGAATCGGA 1305
Db 1263 ATTCATTAACCATTTCACTTGGATTTGATGAGTGAATCTCGCTTCAATGAGTGA 1322
OY 1306 ATTCGTAATTAATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1365
Db 1323 ATTCGTAATTAATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1382
OY 1366 GCGGCTGACACCATGAGGAGTGGTTTCAACGAAAGAGTATCTAACGTTAAGGAGG 1425
Db 1383 GCGGCTGACACCATGAGGAGTGGTTTCAACGAAAGTGTGAGTAAACGTTAAGGAGT 1442
OY 1426 GCGTTCACGAGTGAATTCATGACTGGGGTG 1457
Db 1443 AGCCGCTTAAGTGGGACAGATGATGGGGTG 1474
```

```
RESULT 42
US-11-055-637-78
; Sequence 78, Application US/11055637
; Publication No. US20050260619N1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: DUBOIS, Jason
; APPLICANT: EDGE, Tom
; APPLICANT: MASSON, Luc
; APPLICANT: TREVORS, Jack T.
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
; FILE REFERENCE: 2139-33US
; CURRENT APPLICATION NUMBER: US/11/055,637
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,288
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Probe for DNA array
US-11-055-637-78
```

```
Query Match 57.6%; Score 838.8; DB 17; Length 1507;
Best Local Similarity 76.1%; Pred. No. 3.1e-250;
Matches 117; Conservative 0; Mismatches 332; Indels 19; Gaps 6;

OY 5 AACGTGGCGGCGAGCTTTTACACATGCAAGTCCGACGACGACGAGTGTGCAATCTGG 64
Db 1 AACGTGGCGGCGGCTTAAATACATGCAAGTCCGACGACGACGAGTGTGCAATCTGG 56
OY 65 TGGCAGTGGCGGAGCGGCTGAATGATCGG-AACTATCCAGAAAGGCGGTAACG 123
Db 57 CAATTAGCGGCGGAGCGGCTGAATGATGATCGGCGCAACTGCTTAAAGATCGGTAAC 116
OY 124 CATGAAAGATGCTTAATCCGATATCTTAAAGAGAAAG-CAGGAGATGAAAGA 182
Db 117 TCGGAAACCGGAGTAAATCCGATATGTTCTTCTCGCATGAGAAATGAAAGA 176
OY 183 C-----CTTGGCGTTTGGAGCGCGCATGCTGATTAAGTTAGTTGGGTAAAG 234
```

Db 177 CGGCTCGGCTGTCACTTAATGATGGCCCGCGCATTAGCTAGTTGGTGAAGTATG 236  
Qy 235 GCCTACCAAGCGCAGATCAGTATGCTGTGAGAGACGACGACCACTGGGACTGA 294  
Db 237 GCTTACCAAGCGCAGATGCTGATGCGGACCTGAGAGGATGATCGGCCACTGGGACTGA 296  
Qy 295 GACACGGCCGACGATCTCTACGCGGAGGAGCAGATGGGAAATTTTGAACAATGGCGGACG 354  
Db 297 GACACGGCCGACGATCTCTACGCGGAGGAGCAGATGGGAAATCTTCGCAATGGAAGT 356  
Qy 355 CTGATCCAGCAATGCGCGGTGATGAGAAAGCCCTTGGGTTGTAAGCTCTTCACTG 414  
Db 357 CTGACGAGCAAGCGCCCGGTGAACGATGAAGGCTTTCGGTCTGTAAGGTTCTGTTTGA 416  
Qy 415 AGAAGAAAGGTTACGCTTAATATCGTGAACCCATGACGCTATGACAGAAAGCACCG 474  
Db 417 GAAAGAAAGGTTACGCTTAATATCGTGAACCCATGACGCTATGACAGAAAGCACCG 476  
Qy 475 GCTAACTACGTGCGACGACCGCGGTAAATGTAAGGTGCAAGCGTTAATCGAATTACT 534  
Db 477 GCTAACTACGTGCGACGACCGCGGTAAATGTAAGGTGCAAGCGTTAATCGAATTACT 536  
Qy 535 GGGCGTAAAGGGTCCGACGCGCGCTTGTAGTCAATGTGAATCCCGCGCTTAATCT 594  
Db 537 GGGCGTAAAGGGTCCGACGCGCGCTTGTAGTCAATGTGAATCCCGCGCTTAATCT 596  
Qy 595 GGGAAATTCGCTTGAATCTACAAAGCTAGATGTGGAGAGGAGGTGAATTCATGTG 654  
Db 597 GGGAGGCTCATTTGAAATCTGGGAACTTGAATGTGAAGAGAAAGTGAATTCAGTGT 656  
Qy 655 TAGCAGTAAATGCGTGAAGATATGAAACAATCATGATGGGAAAGGACCTCTGGGTT 714  
Db 657 TAGCAGTAAATGCGTGAAGATATGAAACAATCATGATGGGAAAGGACCTCTGGGTT 716  
Qy 715 AACATGACGCTCATGACGAAAGCGTGGGAGCAACAGATTTAGTACCTCTGTAGTC 774  
Db 717 ATAATCACTAGTGGGCGGAAAGCGTGGGAGCAACAGATTTAGTACCTCTGTAGTC 776  
Qy 775 CAGGCGCTAAACGATGTCATAGTGTGGGCTTATTAAGC--TTGGTAAACAACTA 832  
Db 777 CAGGCGCTAAACGATGTCATAGTGTGGGCTTATTAAGC--TTGGTAAACAACTA 836  
Qy 833 ACGCGTAAATGTCACCGCTGGGAGTACGCTGCGCAAGATTTAAACTCAAGAAATGAC 892  
Db 837 ACGCATTAAGCACTCCGCTGGGAGTACGCGCCGCAAGGCTGAATCAAGAAATGAC 896  
Qy 893 GGGGACCGGCAAGCGGTGATTAATGTGATTAATTTGATGCAACCGGAAACCTTAC 952  
Db 897 GGGGACCGGCAAGCGGTGATTAATGTGATTAATTTGATGCAACCGGAAACCTTAC 956  
Qy 953 CTACCTTGAATGAGGAAATTTCTAAGATAG--ATTAGTCTTGGGAAACGCTAA 1009  
Db 957 CAGGCTTGAATCTCTGACATCTCTAAGATTAAGACGTTCCCTTGGGAAACAGAT 1016  
Qy 1010 CACAGTGTCTGATGCTGTGTCAGCTGCTGTGAGATGTTGGTTAAATGCTCCGCA 1069  
Db 1017 GACAGTGTCTGATGCTGTGTCAGCTGCTGTGAGATGTTGGTTAAATGCTCCGCA 1076  
Qy 1070 CGAGCGCAACCTTGTCTAATTAATGCAATCTTTGGTGGGCACTTTAATGAGACTGCG 1129  
Db 1077 CGAGCGCAACCTTGTCTAATTAATGCAATCTTTGGTGGGCACTTTAATGAGACTGCG 1136  
Qy 1130 GTGCAAAACCGGAGGAGTGGGAGTGAAGTCAAGTCTCTAATGCGCTTATGGGATGGG 1189  
Db 1137 GTGTAACACCGGAGGAGTGGGAGTGAAGTCAATGCTCAATGCTTATTAATGAGACTGGG 1196  
Qy 1190 TTCAACGTAATTAATTAATGCGGCTACAGAGGTTGCGCAACCGCGAGGAGGAGCTTAATCT 1249  
Db 1197 TACACACTGCTACATGATGATGATTAACAAAGCTGCAAAACCGGAGGAGTAAAGCAATCT 1256  
Qy 1250 CAGAAAGCGCTGTAGTCCGAGTGGAGTCTGCAATCTGCACTCGGTGAATGCGAATCG 1309

Db 1257 CATTAAGCATTCTCAGTTCCGATTCAGGCTGCAACTGCGCTGATGAAGCCGGAATCG 1316  
Qy 1310 CTATGTAATCGCGGATCAGATGTCGCGGTGAATACCTTCCGGGTCTTGTACACACGCC 1369  
Db 1317 CTATGTAATCGCGGATCAGATGTCGCGGTGAATACCTTCCGGGTCTTGTACACACGCC 1376  
Qy 1370 CGTCAACACATGGGAGTGGGTTTACACAGAGGAGTACTTAACGTAAGAGAGGCGCT 1429  
Db 1377 CGTCAACACATGGGAGTGGGTTTACACAGAGGAGTACTTAACGTAAGAGAGGCGCT 1436  
Qy 1430 TGCCAGGTGAGATTCATGACTGGGATG 1457  
Db 1437 GCCTAAGTGGGACAGATGATGGGATG 1464

RESULT 43  
US-11-055-637-69  
; Sequence 69, Application US/11055637  
; Publication No. US2005026019A1  
; GENERAL INFORMATION:  
; APPLICANT: BROUSSEAU, Roland  
; APPLICANT: DUBOIS, Jason  
; APPLICANT: EDGE, Tom  
; APPLICANT: MASSON, Luc  
; APPLICANT: TREVORS, Jack T.  
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND  
; FILE REFERENCE: 2139-33US  
; CURRENT APPLICATION NUMBER: US/11/055,637  
; CURRENT FILING DATE: 2005-02-11  
; PRIOR APPLICATION NUMBER: US 60/543,288  
; PRIOR FILING DATE: 2004-02-11  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 69  
; LENGTH: 1486  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Probe for DNA array  
US-11-055-637-69

Query Match 57.5%; Score 838; DB 17; Length 1486;  
Best Local Similarity 76.8%; Pred. No. 5.5e-250;  
Matches 1130; Conservative 0; Mismatches 325; Indels 17; Gaps 8;

Qy 3 TGAACGCTGCGGATCTTTACATGCAAGTGAACGCGACGACGATGCTTGCATCT 62  
Db 3 TGAACGCTGCGGATCTTTACATGCAAGTGAACGCGACGACGATGCTTGCATCT 62  
Qy 63 GGTG--CGAGTGGCGGACGCGGTGATGATGATCGG--AACGTAATCCAGAAAGGGGGT 119  
Db 63 TATGAAGTGAAGCGGCGGACGCGGTGATGATGATGATGATGATGATGATGATGATGAT 122  
Qy 120 AACGATCAAGAAATGCTTAATACCGCAT--ATATCTTAAGAGAGAAAGAGGATGCA 178  
Db 123 AACTCCGGGAAACCGGCGCTTAATACCGCATTAATTTGAACCCATGCTTGAATTTGA 182  
Qy 179 AAGAC-----CTTGCCTTTTGAAGCGCGCATGCTGATTAAGTATGATGATGATGATGAT 230  
Db 183 AAGCGGCTTCGCGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 242  
Qy 231 AAGGCTTACCAAGGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 290  
Db 243 AAGGCTTACCAAGGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302  
Qy 291 CTGAGACAGCGCCGACGCTCTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 350  
Db 303 CTGAGACAGCGCCGACGCTCTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362  
Qy 351 AAGCTGATTCAGCAATGCGCGGTGATGAAAGGCTTCCGGTGTGTAAGCTCTTCA 410  
Db 363 AAGCTGATTCAGCAATGCGCGGTGATGAAAGGCTTCCGGTGTGTAAGCTCTTCA 422

411 GTGAGAGAGAGGTTACGGTA-AATTAATCTGACCCCATGACGGTATCGACAGAAAG 469  
423 TAGAGAGAGAGAGGTTAGTAAGTGGACCTTGAACGGTACTTAACAGAAAG 482  
470 CACCGGCTAATCACTGAGGAGCGCGGGTAAATAGTGGGGGCAAGGGTTAATGGAA 529  
483 CACGGCTAATCACTGAGGAGCGCGGGTAAATAGTGGGGGCAAGGGTTAATGGAA 542  
530 TTAAGGGGCTAAGAGGGTGGAGCGCGGGTAAATAGTGGGGGCAAGGGTTAATGGAA 589  
543 TTAATGGGGTAAAGGGGCGGAGGGTGGTCTTAAAGTGTGTAAGAGCCACGGCTC 602  
590 AACCTGGAAATGCGTTTGAATCAAAAGTGAATGAGTGGGAGAGAGAGTGGAAATTC 649  
603 AACCGTGAAGGGTCAATGGAGAACTTGAAGTGAAGAGAGAGAGAGTGGAAATTC 662  
650 ATGTGAGAGAGTAATGGGTAGAGATATGAGAGAAATCATGATGGGAGAGAGCGCTCT 709  
663 ATGTGAGAGAGTAATGGGTAGAGATATGAGAGAAATCATGATGGGAGAGAGCGCTCT 722  
710 GGGTAACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 769  
723 GGTCTTAATGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 782  
770 TAGTCAACGCTTAAAGATGTCACTAGTTGGGCTTATTAAGC--TTGGTAAACA 827  
783 TAGTCAACGCTTAAAGATGTCACTAGTTGGGCTTATTAAGC--TTGGTAAACA 842  
828 AGTAAAGGCTGAAGTGAAGCGGCTGGGAGTACGGGCAAGATTAAATCAAGAGAA 887  
843 AGTAAAGGCTGAAGTGAAGCGGCTGGGAGTACGGGCAAGATTAAATCAAGAGAA 902  
888 TTGACGGGAGCGGCAAGCGGCTGATTAATGAGTAAATTCATGACCAAGGAGAAAC 947  
903 TTGACGGGAGCGGCAAGCGGCTGATTAATGAGTAAATTCATGACCAAGGAGAAAC 962  
948 CTTAACCTTGAACCTGATTAAGGAAATTTCTGAGATA-GATTAGTCTTGGGAGAGC 1006  
963 CTTAACCTTGAACCTGATTAAGGAAATTTCTGAGATA-GATTAGTCTTGGGAGAGC 1022  
1007 TAAACAGAGTGCATGAGTGCATGAGTGCATGAGTGCATGAGTGCATGAGTGCATG 1066  
1023 AGTAAAGGCTGAAGTGAAGCGGCTGGGAGTACGGGCAAGATTAAATCAAGAGAA 1082  
1067 CAAAGAGCGGCAAGCGGCTGATTAATGAGTAAATTCATGACCAAGGAGAAAC 1126  
1083 CAAAGAGCGGCAAGCGGCTGATTAATGAGTAAATTCATGACCAAGGAGAAAC 1142  
1127 CCGGTGAACAAAGGAGAGAGTGGGAGTAACTCAAGTCTTCAATGGGCTTATGGGTAG 1186  
1143 CCGGTGAACAAAGGAGAGAGTGGGAGTAACTCAAGTCTTCAATGGGCTTATGGGTAG 1202  
1187 GGGTGAACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1246  
1203 GGGTGAACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1262  
1247 TCTCAGAAAGCGGCTGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1306  
1263 TCTCAGAAAGCGGCTGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1322  
1307 TGGCTAATTAAGCGGATGACATGATGCGGATGACATGATGCGGATGACATGATGCGG 1366  
1323 TGGCTAATTAAGCGGATGACATGATGCGGATGACATGATGCGGATGACATGATGCGG 1382  
1367 GCGGTGAACAAAGGAGAGAGTGGGAGTAACTCAAGTCTTCAATGGGCTTATGGGTAG 1425  
1383 GCGGTGAACAAAGGAGAGAGTGGGAGTAACTCAAGTCTTCAATGGGCTTATGGGTAG 1442  
1426 GCGGTGAACAAAGGAGAGAGTGGGAGTAACTCAAGTCTTCAATGGGCTTATGGGTAG 1457  
1443 AGCGGCTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1474

RESULT 44  
US-10-793-626-3356/c  
; Sequence 3356, Application US/107933626  
; Publication No. US2005025478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3356  
; LENGTH: 3169  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-3356  
Query Match 57.3%; Score 835.4; DB 10; Length 3169;  
Best Local Similarity 75.4%; Pred. No. 4.3e-249;  
Matches 1109; Conservative 0; Mismatches 346; Indels 16; Gaps 5;  
3 TGAAGCGTGGCGGCAATGCTTTACATGCAAGTGCAGACGCGACAGCTTGCATCT 62  
2817 TGAAGCGTGGCGGCGGCTTAAATGATGCAAGTGCAGACGCGACAGCTTGCATCT 2758  
63 GGTGGCG--AGTGGCGAGCGGTGATGATGATCGG-AAAGTATCCAGAGAGGGGGGT 119  
2757 TCTGACGTGAGCGGCGAGCGGTGATGATGATGATGATGATGATGATGATGATGATG 2698  
120 AACGATCGAAGATGCTTAATACCGCAT------TACTTAAGAGAGAGAGAGAGAG 171  
2697 AACCTGGAGAAACCGAGCTTAATACCGCATTAATTAATTAATTAATTAATTAATTAAT 2638  
172 GGAATGAAAGACCTTGGCGCTTTGAGCGGCGGATGCTGATTAAGTGAAGTGGTGGTGA 231  
2637 AAGACGGTTTCTGCTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2578  
232 AAGGCTTACCAAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 291  
2577 ACGGCTTACCAAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2518  
292 TGAGACAGCGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351  
2517 TGAGACAGCGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2458  
352 AGCTGATCAGCAATGCGGCGTGAAGTGAAGAGCGCTTGGGTTGTAAGCTCTTTCAG 411  
2457 AGCTGATCAGCAATGCGGCGTGAAGTGAAGAGCGCTTGGGTTGTAAGCTCTTTCAG 2398  
412 TGAAGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 471  
2397 TGAAGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2338  
472 CCGGCTTACCAAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531  
2337 ACGGCTTACCAAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2278  
532 ACTGGAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 591  
2277 ATTGGCGTGAAGAGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2218  
592 CCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 651  
2217 CCGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2158  
652 GTGTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 711

Db 2157 GTGTAGCGGTGAATATGCGAGATATGAGAGAACACAGTGGGGAAGGCACTTTCTGG 2098  
Qy 712 GTTAAACATGACGCTCATGACAGAAACGCTGGGAGCAACAGATTAGATACCTGGTA 771  
Db 2097 TCTGTAACTACACGCTGATGTGCGAAACGCTGGGAGTCAACAGATTAGATACCTGGTA 2038  
Qy 772 GTTCACGCGCTTAAACGATGTCATCTAGTTGTT--GGGCTTATTAGGCTTGTGTAACGAG 829  
Db 2037 GTCCACGCGCTTAAACGATGATGCTAGTGTAGGGGGTTCGCGCCCTTAATGCTGTCAG 1978  
Qy 830 CTAAACGCTGAAGTGAACCCCTGGGGAGTAACGCTGCAAGATTAAACTCAAGAAAT 889  
Db 1977 CTAAACGATTAAAGACTCCGCTGGGAGTAACGCTGCAAGATTAAACTCAAGAAAT 1918  
Qy 890 GACGGGACCCGCAAGCGGTGATATGATGATTATTCGATGCAACCGCAAAACCT 949  
Db 1917 GACGGGACCCGCAAGCGGTGAGCATGTGTTAAATTCGAAGCAACCGCAAAACCT 1858  
Qy 950 TACCTACCTTGAACATGATGCAATTTTCTAGAGATTAATTAAGTGC--TTCCGGAACGC 1006  
Db 1857 TACCAATCTTGAACATCTCTGACCCCTCTAGAGATTAAGTTCCTCCCTTCGGGGAGCAG 1798  
Qy 1007 TAAACAGGTGCTGATGCTGTGTGTCAGCTCGTGTGTCGATGATGTTGGTTAGTCCG 1066  
Db 1797 AGTAAACGATGTGTCATGTGTGTGTGTCAGCTCGTGTGTCGATGATGTTGGTTAGTCCG 1738  
Qy 1067 CAACGACGCGAACCCTTGTCTATTAATTGGCATCATTTGGTGGGCACTTTATGAGACTG 1126  
Db 1737 CAACGACGCGAACCCTTAAAGCTTAAGTTAGTTCATTAAGTTGGGCACTTAAAGTTAGCTG 1678  
Qy 1127 CCGGTGACAAACCGAGAGAGGTGGGAGTAAGTCAAGTCTCTCATGGCCCTTATGAGTAC 1186  
Db 1677 CCGGTGACAAACCGAGAGAGGTGGGAGTGAAGTCAATCATATGCCCCCTTATGATTTG 1618  
Qy 1187 GCGTTCACAGTATTAATCAATGCGCGCTACAGAGGCTTCCCAACCCCGAGGGGAGCTAA 1246  
Db 1617 GCGTTCACAGTATTAATCAATGAGCAATTAACAAAGGCGAGCGAACCAGGCTCAAGCA 1558  
Qy 1247 TCTCAGAAAGCGCGTGTGATGTCGCGATCGAGTCTGCACTGCACTCCGTAAGTCCGAA 1306  
Db 1557 TCCCATTAAGTGTTCATGATTCGATTTGATGTCTGCACTGCACTAATTAAGTCCGAA 1498  
Qy 1307 TCGCTAGTAATCGCGATGACAGATGTCGCGGTGAATAGTTCCCGGCTTGTACACACC 1366  
Db 1497 TCGCTAGTAATCGTAATGATCAGCATGCTACGCTGAATAGTTCCCGGCTTGTACACACC 1438  
Qy 1367 GCGCGTACACAGTGGAGTGGGTTTCAACGAAAGCAAGTATCTAAACCTGAAGAGGCG 1426  
Db 1437 GCGCGTACACAGTGGAGTGGTGTAAACACCCGAAGCCGCTGAGTAACTTTGAGAGCTA 1378  
Qy 1427 GCTTGCAACGCTGAGATTCAATGACTGGGGTG 1457  
Db 1377 GCCGTGCAAGGTGGAGCAAAATGATTTGGGGTG 1347

## RESULT 45

US-10-793-626-3905  
; Sequence 3905, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; PRIOR FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3905  
; LENGTH: 3308  
; TYPE: DNA

ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3905

Query Match 57.3%; Score 835.4; DB 10; Length 3308;  
Best Local Similarity 75.4%; Pred. No. 4,3e-249;  
Matches 1109; Conservative 0; Mismatches 346; Indels 16; Gaps 5;

Qy 3 TGAACCTGCGCGCATGCTTTACATGCAATGCAACGCGACACGCGATGCTTGCATCT 62  
Db 235 TGAACCGTGGCGCGCATGCTTTACATGCAATGCAACGCGACACGCGATGCTTGCATCT 294  
Qy 63 GGTGGG--AGTGGCGAGCGGCTGATTAATGATTCGG--AACGATCCAGAAAGGGGGGT 119  
Db 295 TCTGACGTTTACCGCGGACGCGGTGATTAACGATGATTAACCTTAAAGACTGGGAT 354  
Qy 120 AACGATCCAGAAATGCTAATACCGCATAT-----TACTTAAGAGAGAAAGCAGG 171  
Db 355 AACTTGGGGAACCGGAGCTTAATACCGATTAATATTAAGAACCGCATGTTCAATAGTGA 414  
Qy 172 GATTCGAAGAACCCTTGGCTTTTGAAGCGGCGCATGTCGATTAGCTTGTGGGGTA 231  
Db 415 AAGACGGTTTGTGTCATTTATAGATGATCCGCGCATTAAGCTTATGTTAGTAAGTA 474  
Qy 232 AAGGCTTACCAAGCGGACGATCAATGATGTTGTCTGAGAGGACGACACGACCTGGGAC 291  
Db 475 ACGGCTTACCAAGCGGACGATGCTGATGCGCACTGAAAGGGTATATCGGCCACATGGAC 534  
Qy 292 TGAACAACGCGCCGACATCTCTACGAGGAGCAGAGCTGGGGAATTTTGAACAATGGGCGCA 351  
Db 535 TGAACAACGCTTCCAGATCTCTACGAGGAGCAGAGCTGAGGAATCTTCCGCAATGGGCGAA 594  
Qy 352 AGCCTGATCCAGAAATGCTGATGATGAGAAAGGCTTGGGTTGTAAGCTTTTCAAG 411  
Db 595 AGCCTGACGAGAGCAAGCCGCGTGAAGTGAAGAGTCTTGGATCGTAAACCTCTGTAT 654  
Qy 412 TCGAAGAAAGGTTACGTTAATATCGTGAACCTTACCGTATGATGACAGAAAGCA 471  
Db 655 TAGGAAAGAAACAATGATGATTAATCTATCAGTCTTACCGTATGATGACAAAGCC 714  
Qy 472 CCGGCTTACATGCGGACGAGCGCGGTTAATCGTAGGCTGCAAGGCTTAATCGAAT 531  
Db 715 ACGGCTTACATGCGGACGAGCGCGGTTAATCGTAGGCTGCAAGGCTTAATCGAAT 774  
Qy 532 ACTGGGCTTAAAGGCTGCGAGGCGGCTTGTATAGTGAATGTAATCCCGGGCTTAA 591  
Db 775 ATTGGGCTTAAAGGCGGCTGAGGCGGTTTAAAGTCTGATGTAAGCCACGCGCTCAA 834  
Qy 592 CTTGGGAATGCGTTTAAACTTACAAAGCTAAGATGTCAGAGGAGGAGTTCAT 651  
Db 835 CCGTGAAGGCTTATTTGAAACTGGAATACTGATGTCAGAGAGGAAAGTGAATTCAT 894  
Qy 652 GTTGAAGATGAATATGCTGATGATGATGAAGAACATGATGGGAAGGCGCTCTCG 711  
Db 895 GTTGAAGATGAATATGCTGATGATGATGAAGAACCAAGTGGGAGGCGACTTTCTGG 954  
Qy 712 GTTAAACATGACGCTCATGACAGAAACGCTGGGAGCAACAGATTAGATACCTGGTA 771  
Db 955 TCTGTAACTACACGCTGATGTGCGAAACGCTGGGATTAACAGATTAGATACCTGGTA 1014  
Qy 772 GTTCACGCGCTTAAACGATGTCATCTAGTTGTT--GGGCTTATTAGGCTTGTGAACGAG 829  
Db 1015 GTTCACGCGCTTAAACGATGATGCTAGTGTAGGGGGTTCGCGCCCTTAATGCTGAG 1074  
Qy 830 CTAAACGCTGAAGTGAACCCCTGGGAGTAACGCTGCAAGATTAAACTCAAGAAAT 889  
Db 1075 CTAAACGATTAAAGACTCCGCTGGGAGTAACGCGCAAGGTTGAACTCAAGAAAT 1134  
Qy 890 GACGGGACCCGCAAGCGGTGATTAATGATTAATTCGATGCAACGCGAAACCT 949  
Db 1135 GACGGGACCCGCAAGCGGTGAGCATGTGTTAATTCGAGCAACGCGAAAGAACT 1194

QY 950 TACCTACCTTGACATGTAGCCGAAATTTTCTAGAGATAGATTAATGTC---TTGGGGAACCC 1006  
 |||||  
 DB 1195 TACCAATCTTGACATCTCTGACCCCTTAGAGATAGATTTTCCCTTCGGGGGACAG 1254  
 QY 1007 TAAACAGGTGTCATGCGTGTGTCAGCTGTCGTAGATGTTGGGTTAAGTCCCG 1066  
 |||||  
 DB 1255 AGTAAACAGGTGTCATGCGTGTGTCAGCTGTCGTAGATGTTGGGTTAAGTCCCG 1314  
 QY 1067 CAACAGCCGCAACCTTGTCTAATTAATTCATATTGGTGGGCACTTAAATGAGACTG 1126  
 |||||  
 DB 1315 CAACAGCCGCAACCTTAAAGTTAGTTCATTAATTAATGAGGCACTTAAAGTACAG 1374  
 QY 1127 CCGGTGACAAACCGAGAGAGAGTGGGAGTGAAGTCAAGTCTGATGCGCTTATGGGTAG 1186  
 |||||  
 DB 1375 CCGGTGACAAACCGAGAGAGAGTGGGAGTGAAGTCAATCATCATCCCTTATGATTTG 1434  
 QY 1187 GGCTTACACGTAAATCAATGCGGCGTACAGAGGTTGCAACCCGCGAGGGGAGCTAA 1246  
 |||||  
 DB 1435 GGCTTACACGTGCTCAATGCAATACAAAGGCGAGAAACCGGAGGTCAAGCAAA 1494  
 QY 1247 TCTCAGAAAGCGGTGTAGTCCGATTCGAGTCTGCAACTGACTCCGTGAAGTGGAA 1306  
 |||||  
 DB 1495 TCCCAATAAGTTGTCTCAGTTGCGATTTGATGCTGCAACTGACTATATGAAAGTGGAA 1554  
 QY 1307 TCGCTAGTAATCGCGGATCAGCATGTCCGCTGAATACGTTCCCGGCTTGTACACACC 1366  
 |||||  
 DB 1555 TCGCTAGTAATCGTAATCAGCATGTCTACGCTGAATACGTTCCCGGCTTGTACACACC 1614  
 QY 1367 GCCCGCACACCATGAGGAGTGGGTTTACAGAAAGAGAGTGTCTAAACGTAAGAGAGGC 1426  
 |||||  
 DB 1615 GCCCGCACACCATGAGAGTGGGTTTAAACCCGAAAGCCGCTGAGTAAACATTTGAGACTA 1674  
 QY 1427 GCTTGCACGCTGAGATTCATGACTGGGCTG 1457  
 |||||  
 DB 1675 GCCGTGAAGGTGGACAATGATTTGGGGTGG 1705

RESULT 46  
 US-10-793-626-4187/c  
 ; Sequence 4187, Application US/10793626  
 ; Publication No. US20050255478A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIMBERLY, WILLIAM JOHN  
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 ; FILE REFERENCE: PU348005  
 ; CURRENT APPLICATION NUMBER: US/10/793,626  
 ; CURRENT FILING DATE: 2004-03-04  
 ; PRIOR APPLICATION NUMBER: 60/164,258  
 ; PRIOR FILING DATE: 1999-11-09  
 ; NUMBER OF SEQ ID NOS: 4472  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 4187  
 ; LENGTH: 3657  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 ; OTHER INFORMATION: nucleic acid sequence  
 US-10-793-626-4187

Query Match 57.3%; Score 835.4; DB 10; Length 3657;  
 Best Local Similarity 75.4%; Pred. No. 4,4e-249;  
 Matches 1109; Conservative 0; Mismatches 346; Indels 16; Gaps 5;  
 QY 3 TGAACGCTGGCGGATGCTTACACATGCAAGTGAACGCGACAGCGATGCTTGCATCT 62  
 |||||  
 DB 1625 TGAACGCTGGCGGCTGCTTAATATCATGCAAGTGAACGCGACAGAGAGCTTGTCTC 1566  
 |||||  
 QY 63 GGTGGCG--AGTGGCGGACCGGTGAGTAAATGCAATCGG-AAAGTATCCAGAAAGAGGGGCT 119  
 |||||  
 DB 1565 TCTGAGCTTAAGCGCGGAGCGGTGAGTAAACCGTGAATTAACCTAATTAAGACTGGGAT 1506

QY 120 AACGATCGAAAAGATGTGCTAATACCGGATA-----TACTTAAGAGAGAAAGCAG 171  
 |||||  
 DB 1505 AACCTCGGAAAACCGGACTTAATACCGATATAATATTGAACCGGATGTTCAATATGGA 1446  
 QY 172 GGATGAAAAGACTTGGCGCTTTTGGAGCGGCCGATGTCTGATTAGCTTGTGGGTGTA 231  
 |||||  
 DB 1445 AAGACGGTTTGTCTGTCACTTAATAGATGATCCGCGCGCAATTAAGCTTGTGTGAAGTGA 1386  
 QY 222 AAGGCTTACCAAGGAGGAGATCAAGTAAATGATCTGAGAGAGAGACACGACCACTAGGAGC 291  
 |||||  
 DB 1385 AAGGCTTACCAAGGAGGAGATCAAGTAAATGATCTGAGAGAGAGATCAAGGAGGAGGAGC 1326  
 QY 222 TGAACAGCGCCAGACTCTTACGAGGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 351  
 |||||  
 DB 1325 TGAACAGCGCTTCAAGCTCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266  
 QY 352 AACCTGATCCAGCAATGCGCGGTGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 411  
 |||||  
 DB 1265 AACCTGATCCAGCAATGCGCGGTGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206  
 QY 412 TGAACAGGAAAAGGTTACGCTAAATATCTGATACCCATGAGCGGTATCCGACAGAGAGCA 471  
 |||||  
 DB 1205 TGAACAGGAAAAGGTTACGCTAAATATCTGATACCCATGAGCGGTATCCGATATCCGAAAT 1146  
 QY 472 CCGGCTAATACGTCAGCAGACCGCGGTATATACGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 531  
 |||||  
 DB 1145 AAGGCTTACGTCAGCAGACCGCGGTATATACGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1086  
 QY 532 ACTGGCGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 591  
 |||||  
 DB 1085 ACTGGCGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1026  
 QY 592 CCGTGAATATGCGTTTGAATCAATCAAGCTAAGTGTGACAGAGGAGGAGGAGGAGGAGGAGGAG 651  
 |||||  
 DB 1025 CCGTGAATATGCGTTTGAATCAATCAAGCTAAGTGTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 966  
 QY 652 GTGTAGAGTGAATATGCGTGAAGATATGAGAAATCTTAAAGTGAAGAGGAGGAGGAGGAGGAGGAG 711  
 |||||  
 DB 965 GTGTAGAGTGAATATGCGTGAAGATATGAGAAATCTTAAAGTGAAGAGGAGGAGGAGGAGGAGGAG 906  
 QY 712 GTTAACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTTAGATACCTTGGTA 771  
 |||||  
 DB 905 TCTGTAACTGACGCTCATGACGAAAGCGTGGGAGTCAACAGATTTAGATACCTTGGTA 846  
 QY 772 GTTCAAGCCCTTAAAGATGCTCAATGTTGT--GGGCTTATTAGGCTTGGTAAGCAAG 829  
 |||||  
 DB 845 GTTCAAGCCCTTAAAGATGCTCAATGTTGT--GGGCTTATTAGGCTTGGTAAGCAAG 786  
 QY 830 CTAAAGCGTGAAGTTGACCGCTTGGGAGTACGCTGCAAGATTTAAATCTCAAGAGAT 889  
 |||||  
 DB 785 CTAAAGCGTGAAGTTGACCGCTTGGGAGTACGCTGCAAGATTTAAATCTCAAGAGAT 726  
 QY 890 GAAGGAG 949  
 |||||  
 DB 725 GAAGGAG 666  
 QY 950 TACCTACCTTGACATGTAGCCGAAATTTTCTAGAGATAGATTAATGTC---TTGGGGAACCC 1006  
 |||||  
 DB 665 TACCTACCTTGACATGTAGCCGAAATTTTCTAGAGATAGATTAATGTC---TTGGGGAACCC 606  
 QY 1007 TAAACAGGTGTCATGCGTGTGTCAGCTGTCGTAGATGTTGGGTTAAGTCCCG 1066  
 |||||  
 DB 605 AGTGAACGTTGTCATGCGTGTGTCAGCTGTCGTAGATGTTGGGTTAAGTCCCG 546  
 QY 1067 CAACAGCCGCAACCTTGTCTAATTAATTCATATTGGTGGGCACTTAAATGAGACTG 1126  
 |||||  
 DB 545 CAACAGCCGCAACCTTAAAGTTAGTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 486  
 |||||  
 QY 1127 CCGGTGACAAACCGAGAGAGTGGGAGTGAAGTCAAGTCTGATGCGCTTATGGGTAG 1186  
 |||||  
 DB 485 CCGGTGACAAACCGAGAGAGTGGGAGTGAAGTCAAGTCTGATGCGCTTATGATTTG 426  
 |||||  
 QY 1187 GGCTTACACGTAAATCAATGCGCGTACAGAGGAGGTTGCCAAACCCGCGAGGGGAGACTAA 1246

Db	425	GGCTACACACCTGCTACATGATGGACATATACAAAGGTGACGAAACCGGAGGTCAAGCAA	366
Qy	1247	TCTCAGAAAGCGCGTCGTATGCCGATTCGCACTCGACTCCGTGAATCCGAA	1306
Db	365	TCCCATTAAGTTGTTCTCAGTTCGGATTGTAGTCTGCAACTCGACTATATTAAGCTGAA	306
Qy	1307	TCGCTAGTAATCGCGGATCAGCATGTCCGCGTGAATAGTTCCCGGCTCTTGTACACAC	1366
Db	305	TCGCTAGTAATCGTATGATCAGCATGCTCAACGGTGAATAGTTCCCGGCTCTTGTACACAC	246
Qy	1367	GGCGTCACACATGCGAGTGGGTTTCCACGAAGCAGTATGTAAACCGTAAGGCG	1426
Db	245	GCCGTCACACACGAGATTGTAAACCCGAAAGCGGTGAGTAACATTTTGAAGCTA	186
Qy	1427	GCTTCCACCGGTGAGATTCAATGACTGGGGTG	1457
Db	185	GCCGTCGAAGGTGGGACAAATGATTTGGGGTG	155

```

RESULT 47
US-10-793-626-4460/C
; Sequence 4460, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US-10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4460
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4460

```

Query Match	57.1%;	Score 832.2;	DB 10;	Length 3008;
Best Local Similarity	75.3%;	Pred.No. 4.2e-248;		
Matches 1107; Conservative	0;	Mismatches 348;	Indels 16;	Gaps 57.

Qy	3	TTAAAGCTGGCGGCATGTCTTATACATGCAGATCGAACGGGACAGCGATGCTTCATCT	62
Db	1525	TTAAAGCTGGCGGCCTGCTTAATACATGCAGATCCAGAGAACAGAGAGACTTCCTCC	1466
Qy	63	GGTGGCG--AGTGGCGGACGGGTGAGTAATGCATGG--AACTTATCCAGAAAGGGGGGT	119
Db	1465	TCGTGACGTAAAGCGGCGGACGGGTGAGTAACAATGGATTAACCTATTAACATCGGAT	1406
Qy	120	AACGCATGGAAGATGTGTAAATACCGCAT-----TACTTAAGAGGAAAGACAG	171
Db	1405	AACCTCGGGAACCGGAGCTTAATCCGATTAATATTGAACCGCATGTTCAATATGTA	1346
Qy	172	GGATCGAAAGACCTTGCGCTTTTGAGCGGCGGATGTCTGATTAGCTAAGTTGTGGGTA	231
Db	1345	AAGACGGTTTTCGTGCACTTAATAGTGAATCCGCGCCGCAATTAAGTAGTTGGTAAGTA	1286
Qy	232	AAGGCTTACCAAGGCGACGATCAGTATGTGGTCTGAGAGACGACGACACATGGGAC	291
Db	1285	ACGGCTTACCAAGGCAACGATGTGTAGCCGACTTGAGAGGGTGTATGGCCACATCGAAC	1222
Qy	292	TTGAGACAGCGCCAGACTCTTACGAGGAGCAGCAGTGGGAAATTTTGGACAATTGGCGCA	351
Db	1225	TTGAGACACGCTCCAGACTCTTACGAGGAGCAGCAGTAGGGAATCTTCGCAATTGGCGCA	1166
Qy	352	AGCTGATCCAGCAATGCCGCGTAGTGAAGAAAGCCTTCGGGTTGTAAAGCTCTTTCAG	411

[illegible]



RESULT 48  
US-10-793-626-3955/c  
; Sequence 3955, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patentl Ver. 2.1  
; SEQ ID NO 3955  
; LENGTH: 3253  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3955

Query Match 57.1%; Score 832.2; DB 10; Length 3253;  
Best Local Similarity 75.3%; Pred. No. 4,2e-248;  
Matches 1107; Conservative 0; Mismatches 348; Indels 16; Gaps 5;

3 TGAACGCTGGCGCATCTTTACACATGCAATGCAACGCGACGATGCTTGATCT 62  
1476 TGAACCTGGCGCGCTGCTTATATACATGCAATGCAACGCGACGATGCTTGCTTC 1417  
63 GGTGGCG -AGTGGCGACGGGTGATGATCGG -AACGATCCAGAAAGGGCGT 119  
1416 TCTGACGTTAGCGCGGACGGGTGATGATGCAACGCGATGATGCTTATGAGCTGGAT 1357  
120 AAGCATCCGAATATGCTTATACCGCATATCTCT-----AAGGAGAAAGAGG 171  
1356 AACTTGGGAAACCGGAGCTTATACCGGATATATATTGAAACCGCATGCTTCAATGTA 1297  
172 GGATCGAAAGACCTTGGCGCTTTTGGAGCGCGCATGCTGATTAGCTAGTTGGTGGGTA 231  
1296 AAGACGCTTTTGTCTCTTATATAGATGATCGCGCGCATTAAGTATGTTGAAGTA 1237  
232 AAGGCTTACCAAGGCGACGATCAGTATGCTGAGAGACGACGACCACTGGGAC 291  
1236 ACGGCTTACCAAGGCGACGATGCTGAGCGACCTGAGAGGGTATCGGCCACACTGGAAC 1177  
292 TGAAGACCGGCGACGATCTCTACGGGAGGAGACGATGGGGAATTTTGGACAATGGGCGCA 351  
1176 TGAAGACCGGCTCCAGATCTCTACGGGAGGAGACGATGGGGAATTTTCCGCAATGGGCGAA 1117  
352 AGCCTGATCCAGCAATCCCGCTGAGTGAAGAAAGGCTTCGGGTTGTAAGCTCTTTCAG 411  
1116 AGCCTGACGAGACGACCGCGCTGAGTGAAGAGTCTTCGATCGTAAACTCTGTAT 1057  
412 TCGAAGAAAGAGTTACGGTAAATATCGTGAACCCATGACCGTATGACAGAAAGCA 471  
1056 TAGGAAAGAACTATGTAATGTAATCACTGTTTACCGGTATCTAATCAGAAAGCC 997  
472 CCGGCTTACCTAGTCCAGACGACCGCGGTATATAGTGGGTGGAAGGCTTAATCCGAAT 531  
996 ACGGCTTACCTAGTCCAGACGACCGCGGTATATAGTGGGTGGAAGGCTTAATCCGAAT 937  
532 ACTGGGCTTAAAGGCTGCGAGGCGGCTTGAAGTCAATGTAATCCCGGCTTAA 591  
936 ATTGGGCTTAAAGGCTGCGAGGCGGCTTGAAGTCAATGTAATCCCGGCTTAA 877  
592 CCTGGGAATGCGTTTGAACCTAACAACTAGAGTGGGAGAGGAGGTGGAATTCAT 651  
876 CCGTGGAGGCTCAATTGGAACCTGGAACCTTGAAGTGAAGAGAAAGTGAATTCAT 817

652 GTGTAGCAGTGAATTCGTTAGATATGGAAGAACATCGATGCGAGACGCTTCCTGG 711  
816 GTGTAGCAGTGAATTCGTTAGATATGGAAGAACATCGATGCGAGACGCTTCCTGG 757  
712 GTTAACTGACGCTTATGACGAAAGGCTGGGAGCAACAGATTAATATCCCTGGTA 771  
756 TCTGTAACTGACGCTTATGACGAAAGGCTGGGAGCAACAGATTAATATCCCTGGTA 697  
772 GTTCAAGGCTTAAAGATGCACTAGTTGT--GAGGCTTATTAAGCTTGGTAAAGAA 829  
696 GTTCAAGGCTTAAAGATGCACTAGTTGT--GAGGCTTATTAAGCTTGGTAAAGAA 637  
830 CTAAAGGCTTAAAGATGCACTAGTTGT--GAGGCTTATTAAGCTTGGTAAAGAA 889  
636 CTAAAGGCTTAAAGATGCACTAGTTGT--GAGGCTTATTAAGCTTGGTAAAGAA 577  
890 GACGGGAGACCGCAACAGGCTGATATGATGATTAATTCATGCAACGCGAAACCT 949  
576 GACGGGAGACCGCAACAGGCTGATATGATGATTAATTCATGCAACGCGAAACCT 517  
950 TACCTAACCTTGAACATGTAAGGAAATTTCTAGAGATGATGATGCT--TTGGGAAAGC 1006  
516 TACCAATTTGACATCTCTGACCCCTTGAAGATGATGATGCT--TTGGGAAAGC 457  
1007 TAAACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066  
456 AAGTACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 397  
1067 CAACGAGCGCAACCTTGTCTATTAATGCTCATATTTGGTGGCACTTAAATGAGACTG 1126  
396 CAACGAGCGCAACCTTGTCTATTAATGCTCATATTTGGTGGCACTTAAATGAGACTG 337  
1127 CCGGTGACAAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1186  
336 CCGGTGACAAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 277  
1187 GCGTTCACACGTAATACATGCGGCTACAGAGGCTTGGCAACCGCGAGGAGGAGGAGGAG 1246  
276 GCGTTCACACGTAATACATGCGGCTACAGAGGCTTGGCAACCGCGAGGAGGAGGAGGAG 217  
1247 TCTCAGAAAGCGGCTGATGATGCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1306  
216 TCCATTAAGTGTCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 157  
1307 TCGCTAGTATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366  
156 TCGCTAGTATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 97  
1367 GCCCGTCAACCATGAGAGTGGGTTTCAACAGAGGAGTATCTAACCGTAAAGAGGCG 1426  
96 GCCCGTCAACCATGAGAGTGGGTTTCAACAGAGGAGTATCTAACCGTAAAGAGGCG 37  
1427 GCTTGCACCGTGAATTCATGATGAGGCTG 1457  
36 GCCGTGAAAGTGGGCAAAATGATGAGGCTG 6

RESULT 49  
US-10-793-626-3967/c  
; Sequence 3967, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patentl Ver. 2.1  
; SEQ ID NO 3967  
; LENGTH: 3821



TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3967

Query Match 57.1%; Score 832.2; DB 10; Length 3821;  
Best Local Similarity 75.3%; Pred. No. 4,4e-248;  
Matches 1107; Conservative 0; Mismatches 348; Indels 16; Gaps 5;

3 TGAACGCTGGCGGCGATCTTACACATGCAAGTCAACGCGACGACGATGCTTGACATCT 62  
3624 TGAACGCTGGCGGCGGCTCTTAATACATGCAAGTCAACGCGACGACGATGCTTGACATCT 3565  
63 GGTGGCG--AGTGGCGGACGCGGTGATGATGATCGG-AAAGTATCCAGAAAGAGGCGGT 119  
3564 TCTGACCTTACGCGCGGACGCGGTGATGATGATCGG-AAAGTATCCAGAAAGAGGCGGT 3505  
120 AACGCAATGAAAGATGCTTAATACGCGATG-----TACTTAAGAGAGAAAGACAG 171  
3504 AACCTGGGAAACCGGAGCTTAATACGCGATGATGATGATGATGATGATGATGATGATG 3445  
172 GAATCGAAAGACCTTGGCGCTTTTGGACGCGCGGATGCTGATGATGATGATGATGATG 231  
3444 AAGACGGTCTTGTCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3385  
232 AAGGCTTACCAAGGCGACGATCAGTATGCTGATGATGATGATGATGATGATGATGATGAT 291  
3384 ACGGCTTACCAAGGCGACGATCAGTATGCTGATGATGATGATGATGATGATGATGATGAT 3325  
292 TGAGACACGCGCCAGACTCTCTACGCGAGGACGAGAGGCGGAAATTTGGACATGCGGCG 351  
3324 TGAGACACGCGCCAGACTCTCTACGCGAGGACGAGAGGCGGAAATTTGGACATGCGGCG 3265  
352 AGCTGATCCAGCAATCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 411  
3264 AGCTGATCCAGCAATCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 3205  
412 TCGAGAGAAAGGTTACGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 471  
3204 TAGGAGAGAAAGGTTACGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3145  
472 CCGGCTAATCTAGCTGCGACGACCGCGGTAACTAGTACGCTGCGACGCTTAATCGGAAT 531  
3144 ACGGCTAATCTAGCTGCGACGACCGCGGTAACTAGTACGCTGCGACGCTTAATCGGAAT 3085  
532 ACTGGGCGTAAAGGCGTGGCGAGCGGCTTTGTAAGTCAATGATGATGATGATGATGATG 591  
3084 ATTGGGCGTAAAGGCGTGGCGAGCGGCTTTTAAAGTCAATGATGATGATGATGATGATG 3025  
592 CTTGGGAATGCGTTTGAATTAACAAAGTACAGTGGCGAGAGGAGTGGAAATTCAT 651  
3024 CCGTGGAGGATCATTTGAACTGGAACCTTGAAGTCAAGAGAGGAGTGGAAATTCAT 2965  
652 GTGTAGCAGTAAATGCTAGAGATGATGAGAAAGTCAATGATGATGATGATGATGATGATG 711  
2964 GTGTAGCAGTAAATGCTAGAGATGATGAGAAAGTCAATGATGATGATGATGATGATGATG 2905  
712 GTTAACTAGCAGCTGATGACGAAAGCTGGGGAGCAACAGATGATGATGATGATGATGAT 771  
2904 TCTGTAATGACGCTGATGCGAAAGCTGGGGAGTCAAAACGATGATGATGATGATGATGAT 2845  
772 GTTCAAGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 829  
2844 GTTCAAGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2785  
830 CTAAAGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 889  
2784 CTAAAGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2725  
890 GACGGGAGCCGCGACAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 949

Db 2724 GACGGGAGCCGCGACAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 2665  
Qy 950 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1006  
Db 2664 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2605  
Qy 1007 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066  
Db 2604 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2545  
Qy 1067 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126  
Db 2544 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2485  
Qy 1127 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186  
Db 2484 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2425  
Qy 1187 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246  
Db 2424 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2365  
Qy 1247 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306  
Db 2364 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2305  
Qy 1307 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366  
Db 2304 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2245  
Qy 1367 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426  
Db 2244 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2185  
Qy 1427 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
Db 2184 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2154

RESULT 50  
US-10-953-392-1  
Sequence 1, Application US/10953392  
Publication No. US20050289667A1  
GENERAL INFORMATION:  
APPLICANT: CAMBIA  
TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells  
FILE REFERENCE: CAMBIA 414B  
CURRENT APPLICATION NUMBER: US/10/953,392  
CURRENT FILING DATE: 2004-09-28  
PRIOR APPLICATION NUMBER: US 60/583,426  
PRIOR FILING DATE: 2004-06-28  
NUMBER OF SEQ ID NOS: 83  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1  
LENGTH: 1408  
TYPE: DNA  
ORGANISM: Rhizobium spp. NGR234 (strain ANU240)  
US-10-953-392-1

Query Match 57.1%; Score 832; DB 10; Length 1408;  
Best Local Similarity 77.8%; Pred. No. 4e-248;  
Matches 1087; Conservative 0; Mismatches 285; Indels 46; Gaps 5;

Qy 66 GGCAGTGGCGGACGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 125  
Db 31 GGCAGTGGCGGACGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 90  
Qy 126 TCGAAGATGCTTAATACGCGATATCTTAAAGAGAAAGCGGCGATGAAAGACT 185  
Db 91 GCGAAGCTTGTGATTAATACGCGATATCTTAAAGAGAAAGCGGCGATGAAAGACT 145  
Qy 186 TCGGCTTTTGAAGCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 245

```
Db 146 -----GANTGACCGCGCTTGAGTTAGCTTGGTGGGTAAAGGCTTACCAAG 196
Qy 246 CGACGATCAGTATGCTGTCTGAGAGACGACCGCACTGGGACTGAGACCGGCCCA 305
Db 197 CGACGATCAGTATGCTGTCTGAGAGATGATCAGCCACTGGGACTGAGACCGGCCCA 256
Qy 306 GACTCTTACGGGGGAGGAGCGTGGGGATTTTGGACAAATGGGGCGCAAGCTGTATCCAGA 365
Db 257 AACTCTTACGGGGGAGGAGCGTGGGGATTTTGGACAAATGGGGCGCAAGCTGTATCCAGC 316
Qy 366 ATGCCCGGTGAGTGAAGAAAGGCTTGGGTTGTAAAGCTTTCACTCGAAGAAAG 425
Db 317 ATGCCCGGTGAGTGAAGAAAGGCTTGGGTTGTAAAGCTTTCACTCGAAGAAAGATA--- 373
Qy 426 TTACGGTAAATAATCGTGAACCATGACGGTATGACAGAAAGAACCGGCTTAACTACGT 485
Db 374 -----ATGACGGTTAACCGGAGAAAGAACCCCGGCTTAACTTCGT 411
Qy 486 GCCAGAGCGCGGTTAATAGTGGGTGCAAGGGTTAATCGAATTACTGGGGCTAAAG 545
Db 412 GCCAGAGCGCGGTTAATAGTGGGTGCAAGGGGTAGCGGTGTTGGAATTACTGGCGTTAAAGC 471
Qy 546 GTSCGAGCGCGCTTGTATAGTCAAGTGTAAATCCCGGGCTTAACTTGGGAATTGGCT 605
Db 472 GCACGTAGCGCGCATTTAAGTCAAGGGGTGAATCCCGGGGCTCAACCCCGGAATGGCT 531
Qy 606 TTGAAACTACAAAGCTAAGTGTGGCAGAGGAGGTGAATTCATGTGTAGCATGAA 665
Db 532 TTGATACTGGGTGTCTAAGATCCGGAAGAGTGAATTCGAGTGTAGAGTGAA 591
Qy 666 TGGGTAGAGTATGGAAGAACATCGAATGGGAAGGCGCTCTCGGGTTAACTAGACGC 725
Db 592 TTGATAGATTTTGGAGGAACACAGTGGCGAAGGCGCTCACCTGGTCCGTACTGACGC 651
Qy 726 TCATGACGAAGCGTGGGAGCAACAGATTAAGATCCCTGTAGTCCAGCCCTTAA 785
Db 652 TGAGGTGCGAAGCGTGGGAGCAACAGATTAAGATCCCTGTAGTCCAGCCCTTAA 711
Qy 786 CGATGTCACTAGTGTGGGCTTATTAAGCTTGTGAACGAAGCTTAAACCGTGAAGTTG 845
Db 712 CGATGAATGTTAAGCCGTGGGAGTTTACTGTTCGGGTGGCGCACTAACCATTTAAACAT 771
Qy 846 ACCGCTTGGGAGTACGCTGCGCAAGATTAATCTCAAGAAATTAAGCGGGGACCGGACA 905
Db 772 TCCGCTTGGGAGTACGCTGCGCAAGATTAATCTCAAGAAATTAAGCGGGGACCGGACA 831
Qy 906 AGGGTGAATTAATGATTAATTCATGCAACGCGAAACCTTACTTACCTTGAAT 965
Db 832 AGCGTGAACATGTGTTAATTCGAAGCAACGCGAATCTTACCAAGCTTGAAT 891
Qy 966 GTAGCGAATTTTCTAGAGA-----TAGATTAGTCTTCGGAACGCTTACACAGTGTG 1020
Db 892 CCGGCTTGGGATACGAGAGATGCTTACTTCACTTCGCTGACCGGACAGAGTGTG 951
Qy 1021 CATGCTGTGTGCTAGCTGCTGTGTGTGATGTTGGTTAAGTCCCGCAAGACGCGAC 1080
Db 952 CATGCTGTGTGCTAGCTGCTGTGTGTGATGTTGGTTAAGTCCCGCAAGACGCGAC 1011
Qy 1081 CTGTGATTAATTAATGCACTATTTGGTGGGCACTTAAATGAGACTGCGGGTGAACAAACG 1140
Db 1012 CTGCGCTTATGTTGCGACATTTGGTGGGCACTTAAAGGGGACTGCGGGTGAATGACCG 1071
Qy 1141 -GAGGAAGTGGGATGATGATGCTCAATGCGCTTATAGGTAGGGCTTCAACAGTA 1199
Db 1072 AGAGGAAGTGGGATGATGATGCTCAATGCGCTTATAGGTAGGGCTTCAACAGTA 1131
Qy 1200 ATACATATGCGCGTCAAGAGGTTGGCCACCCGCGAAGGGGAGTAACTTCAAGAAAGCG 1259
Db 1132 CTACATATGTTGTGACAGTGGGCGCAGAGACCGGAGAGTGAAGTGAATCTTCAAAAGC-C 1190
Qy 1260 GTGCTAGTCCGGAATCGAGTCTGCACTCGATCCGGAAGTGGGAATCGTGTAGTATCG 1319
Db 1191 ATCTGATGTGATGCACTCTGCACTCGAGTGCATGAAGTTGAATCGTGTAGTATCG 1250
```

```
Qy 1320 CGATCAGCATGTGCGGGTGAATTAAGTTCCCGGGTCTTGTACACACCGCCGTCAACCA 1379
Db 1251 CAGATCAGCATGTGCGGGTGAATTAAGTTCCCGGGTCTTGTACACACCGCCGTCAACCA 1310
Qy 1380 TGGAGTGGGTTTACACAGAACAGTGTCTTAAACCTTAAGAGGAGCGCTTCCACGTTG 1439
Db 1311 TGGAGTGGGTTTCTACCCGGAAGTGTGCTTAAACCGCAAGAGGAGCTTAAACAGGTA 1370
Qy 1440 AGATTCAATGATCTGGGGTG 1457
Db 1371 GGGTACGACCTGGGGTG 1388

RESULT 51
US-10-954-147-1
; Sequence 1, Application US/10954147
; Publication No. US20050289672A1
; GENERAL INFORMATION:
; APPLICANT: CAMBIA
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells
; FILE REFERENCE: CAMBIA 414
; CURRENT APPLICATION NUMBER: US/10/954,147
; PRIOR FILING DATE: 2004-09-28
; PRIOR APPLICATION NUMBER: 60/583,426
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1
; LENGTH: 1408
; TYPE: DNA
; ORGANISM: Rhizobium spp. NGR234 (strain ANU240)
US-10-954-147-1
```

```
Query Match 57.1%; Score 832; DB 10; Length 1408;
Best Local Similarity 77.8%; Pred. No. 4e-248;
Matches 1087; Conservative 0; Mismatches 265; Indels 46; Gaps 5;
```

```
Qy 66 GGGGATGCGCGGATCGGTGATTAATCGAATCGGAACGTATCCAGAAAGGGGGTAAACGCA 125
Db 31 GGGGATGCGCGGATCGGTGATTAATCGAATCGGAATCTTCTTCTACGAATTAACGCA 90
Qy 126 TCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAACAGGGATCGAAAGACCT 185
Db 91 GGGAACTTGTGCTAATACCGTATGAGCCCTTCGGGGGAAAGATTATTCGGGAAA----- 145
Qy 186 TGGCTTTTGAAGCGCGCATGTCTGATTAGCTAGTTGGTGGGTTAAAGGCTTACCAAG 245
Db 146 -----GANTGACCGCGCTTGAGTTAGCTAATTTGGGTGAAGGCTTACCAAG 196
Qy 246 CGACGATCAGTATGCTGTCTGAGAGACGACCGCACTGGGACTGAGACCGGCCCA 305
Db 197 CGACGATCAGTATGCTGTCTGAGAGATGATCAGCCACTGGGACTGAGACCGGCCCA 256
Qy 306 GACTCTTACGGGGGAGGAGCGTGGGGATTTTGGACAAATGGGGCGCAAGCTGTATCCAGA 365
Db 257 AACTCTTACGGGGGAGGAGCGTGGGGATTTTGGACAAATGGGGCGCAAGCTGTATCCAGC 316
Qy 366 ATGCCCGGTGAGTGAAGAAAGGCTTGGGTTGTAAAGCTTTCACTCGAAGAAAG 425
Db 317 ATGCCCGGTGAGTGAAGAAAGGCTTGGGTTGTAAAGCTTTCACTCGAAGAAAGATA--- 373
Qy 426 TTACGGTAAATAATCGTGAACCATGACGGTATGACAGAAAGAACCGGCTTAACTACGT 485
Db 374 -----ATGACGGTTAACCGGAGAAAGAACCCCGGCTTAACTTCGT 411
Qy 486 GCCAGAGCGCGGTTAATAGTGGGTGCAAGGGTTAATCGAATTACTGGGGCTAAAG 545
Db 412 GCCAGAGCGCGGTTAATAGTGGGTGCAAGGGGTAGCGGTGTTGGAATTACTGGCGTTAAAGC 471
Qy 546 GTSCGAGCGCGCTTGTATAGTCAAGTGTAAATCCCGGGCTTAACTTGGGAATTGGCT 605
Db 472 GCACGTAGCGCGCATTTAAGTCAAGGGGTGAATCCCGGGGCTCAACCCCGGAATGGCT 531
```

Qy	606	TTGAAACCTAACAAAGCTAGAGTGTGGGACAGAGGAAGTGGAAATTCCATGTGTAGCAGCTGAAA	665
Db	532	TTGTACTCTGGGTGCTTAGAGTCCGGAAGAGGTAGTGGAAATTCGAGTGTGAGAGTGA	591
Qy	666	TGCGTAGAGATATGGAAGAACATCGATGCGGAAAGGACGCTCTGGGTTAAACATGACG	725
Db	592	TTGCTAATATTTGGAGAGAACACAGATGGCGGAAGCGGCTCACTGATCCGGTACTGACGC	651
Qy	726	TCATGCACGAAAGCGTGGGAGGCAAAACAGATTAGATACCTTGGTAGTCCACGCCCTTAA	785
Db	652	TGAGGTGCGAAAGCGTGGGAGGCAAAACAGATTAGATACCTTGGTAGTCCACGCCGTTAA	711
Qy	786	CGATGTCAACATGATTGTTGGGCTCTTATTAAGGCTTGGTAGTAAACGAACCTAACGCGTAATTG	845
Db	712	CGATGAATTTTAGCCCTCGGCACTTTACTGTTGCGTGGCGCACTTAACGATTTAAACAT	771
Qy	846	ACCGCTGGGAGATACGGTGCAGAAATTAAACTCAAGAAATTGACGGGAGACCCGCA	905
Db	772	TCGCGCTGGGAGATACGGTGCAGAAATTAAACTCAAGAAATTGACGGGAGACCCGCA	831
Qy	906	AGCGGTGATTAATGTGATTTAATTCGATGCAACGCAAAAACCTTACCTACCTTGACAT	965
Db	832	AGCGGTGAGATGTGATTAAATTCGAAAGCAACGCGCAAACTTACCAAGCTTGACAT	891
Qy	966	GTACGGAATTTTCAGAGA-----TAGATTAGTCTTCGGGAGCGCTAACAGAGTGTG	1020
Db	892	CCCGGTGCGGATACGAGAGATCGATTCCTCAGTTGGGCTGAGACGGAGACAGGTGTG	951
Qy	1021	CATGGCTGTGTCAGCTGCTGTGCTGAGATGTTGGGTTAAGTCCGCAACAGCGCAAC	1080
Db	952	CATGGCTGTGTCAGCTGCTGTGCTGAGATGTTGGGTTAAGTCCGCAACAGCGCAAC	1011
Qy	1081	CTTGTCAATTAATGCCATCATTTGGTGGGCACTTAAAGAACCTGCCGCTGACAAACG	1140
Db	1012	CTGCGCTTAAATGTGCAGCACTTTGGTGGGCACTTAAAGGGAATGCGCGGTATGAGCG	1071
Qy	1141	-GAGGAAGGTGGGATGACGTCAAGTCTCAATGCGCTTATGGGTAGGGCTTCAACGTA	1199
Db	1072	AGAGGAAGTGGGATGACGTCAAGTCTCAATGCGCTTAAAGGGAATGCGCGGTATGAGCG	1131
Qy	1200	ATPACATGCGCGGTACAGAGGGTTCACCAACCCGAGAGGGAGACTAATCTCAGAAAGCG	1259
Db	1132	CTACAATGTGTGTGACAGTGGGACGCGAGACCGGAGGTTCAGGCTTAACTTCAAAAGC-C	1190
Qy	1260	GTCGTAGTCGGATGCGAGTCTGCAACCTCGACTCGTAGTCGGAAATCGCTAGTAATCG	1319
Db	1191	ATCTCAGTTCGAAATGCACTCTGCAACTCGAGTGCATGAAATGGAAATCGCTAGTAATCG	1250
Qy	1320	CGGATACAGCATGTGCGCGGTGATACGTTCCGCGGCTTGTACACACCGCCGTCACCA	1379
Db	1251	CAGATACAGCATGTGCGCGGTGATACGTTCCGCGGCTTGTACACACCGCCGTCACCA	1310
Qy	1380	TGGAGTGGGTTTACCGAAGACAGGTAGTCTAACCGTAAGAGAGGCGCTTGCAACGCT	1439
Db	1311	TGGGAGTGGTCTTACTACCGAAGGTAGTGCCTAACCGAAGAGGCACTAACACGCTA	1370
Qy	1440	AGATTCAATGACTGGGGTG 1457	
Db	1371	GGGTACGCACTGGGGTG 1388	
RESULT 52			
US-10-953-392-15			
; Sequence 15, Application US/10953392			
; Publication No. US2005028967A1			
; GENERAL INFORMATION:			
; APPLICATION: CAMBIA			
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells			
; FILE REFERENCE: CAMBIA 414B			
; CURRENT APPLICATION NUMBER: US/10/953,392			
; CURRENT FILING DATE: 2004-09-28			
; PRIOR APPLICATION NUMBER: US 60/583,426			

Query Match	57.1%	Score 832	DB 10	Length 1433
Best Local Similarity	77.8%	Pred. No. 4e-248		
Matches 1087	Conservative 0	Mismatches 265	Indels 46	Gaps 5
PRIOR FILING DATE: 2004-06-28 NUMBER OF SEQ. ID NOS: 83 SOFTWARE: PatentIn version 3.2 SEQ ID NO 15 LENGTH: 1433 TYPE: DNA ORGANISM: A. tumefaciens EHA105 US-10-953-392-15				
QY	66	GGCGAGTGGCGGACGGGNGAGTATGCAATCGAAGCTATCCAGAAAGAGGGGGTAAACCA	125	
DB	20	GGGAGATGGCGAGACGGGTAGTAAACGGGTGGGAATTAACCACTCTCTGGGAATAGCTCT	79	
QY	126	TCGGAAGATGTCTAATATACCGCATATATCTTAAAGAGAAAGACAGGGGATCGAAAGACT	185	
DB	80	GGGAATCTGGAATTAATATACCGCATACCCCTACGGGG-----GAAAGATT	125	
QY	186	TGCGCTTTTGGAGCGCGCCGATGTCGATTAAGTACTGTTGGTGGGGTAAAGGCTCCAGAG	245	
DB	126	ATTCGGGATGGATGAGCCCGGTTGGATCTATGTTGGTGGGGTAAAGGCTTACCAAG	185	
QY	246	CGACGATCAGATGTTGGTCTGAGAGGACGACACACCACTAGGAGCTGAGACACGGCCCA	305	
DB	186	CGACGATTCATAGCTGTCTGAGAGGATGATACGCCACATTTGGAGCTGAGACACGGCCCA	245	
QY	306	GACTCTCAACGGGAGGCGACGACTGGGGAATTTTGGACAAATGGCGCGACCTGATTCAGCA	365	
DB	246	AACTCTCAACGGGAGGCGACGACTGGGGAATTTTGGACAAATGGCGCGACCTGATTCAGCC	305	
QY	366	ATGCCCGGTAGTGAAGAGGCTTCGGGTGTAAAGCTCTTCACTGAGTGAAGAAAG	425	
DB	306	ATGCCCGGTAGTGAAGAGGCTTCGGGTGTAAAGCTCTTCACTGAGTGAAGAAAG	362	
QY	426	TTACGGTAAATATATCGTACCCATGACGGTATCCACAGAAAGACCGGCTAATCTAGT	485	
DB	363	-----ATGACGGTATGTGGAGAAAGACCGGCTAATCTAGT	400	
QY	486	GCCAGACCGCGGTAAATACGTAGGGGTGCAGAGCTTAAATTCGAAATTAATCTGGCGTTAAAG	545	
DB	401	GCCAGACCGCGGTAAATACGAAAGGGGCTGAGCTTGTTCGAAATTAATCTGGCGTTAAAGC	460	
QY	546	GTCGGCAGACCGGCTCTGTAATCTCAAGTGTAAATCCCGGGCTTAACTGGGAAATTTGCGT	605	
DB	461	GCACTGATGCGGATATTTAAATCTGACGGGTGTAATCCCGCACTCAACTCGGAAATCTGCT	520	
QY	606	TTGAAATCTCAAAAGGTAGATGTGTGACAGAGGAGGTGGAATTCATGTGTAGCACTGAAA	665	
DB	521	TTGATCTAGGATATCTTGAATGTGAAGAGGTAACTGGAATTTCCGAGTGTAGAGGTGAAA	580	
QY	666	TGCGTAGATATGGAAGAACATCGATGGCGAAGGCGACGCTCTGGGTTAAACCTGAGCGC	725	
DB	581	TTTCGTGAATATTCGGAGGAACACAGATGGCGAAGGCGGCTTATCTGCTCATTAATCGAGCGC	640	
QY	726	TCATGACAGAAAGGTGGGGAGCAAAACAGATTTGATATCCCTGTGAATGCACGCGCTTAA	785	
DB	641	TGAGGTGGGAAGCGTGGGGAGCAAAACAGATTTGATATCCCTGTGAATGCACGCGCTTAA	700	
QY	786	CGATGTCAATAGTGTGGGCTTATTAAGGCTTGTGTAACGAAAGCTAAACGCGTGAAGTTG	845	
DB	701	CGATGAATAGTGTGGGCTTGTGTAACGAAAGCTTGTGTAACGAAAGCTTGAAGCAT	760	
QY	846	ACCGGCTGGGAGTACGCTCGCAAGATTTAAATCTCAAGAAATTTGACGGGGACCCGACAC	905	
DB	761	TCCGCTTGGGAGTACGCTCGCAAGATTTAAATCTCAAGAAATTTGACGGGGACCCGACAC	820	
QY	906	AGCGGTGATTAATGTGATTAATTCGATGACGCAAGCGGAAAAACCTTAACCTACCTTGAAT	965	
DB	821	AGCGGTGATTAATGTGATTAATTCGATGACGCAAGCGGAAACCTTAACCAAGCTTTGAAT	880	

QY 966 -----GTACGAATTTCTAGAGATAGATTAGCTTCGGGAACGGTAAACAGAGTCTG 1020  
DB 881 TCGGGGTAATGGGATTTGGAGACATGTCCTTAGTAAAGCTGGCCCCAAGAACAGGTGCTG 940  
QY 1021 CATGGCTGTCAGCTCGTGTCTGATGATGTTGGGTTAAAGTCCGCAACAGCGCAACC 1080  
DB 941 CATGGCTGTCAGCTCGTGTCTGATGATGTTGGGTTAAAGTCCGCAACAGCGCAACC 1000  
QY 1081 CTTGTCAATTAATGGCATCATTTGGTGGGCACTTTAATGAGACTCGCGGTGCAAAACG 1140  
DB 1001 CTCGCCCTTAGTGGCCAGCATTTAGTTGGGCACTTAAGGGGACTGCGCGGTGTAAGCGG 1060  
QY 1141 -GAGGAAGGTGGGATGACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACGTA 1199  
DB 1061 AAGGAAGGTGGGATGACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACGTA 1120  
QY 1200 ATACAATGGCGCGCTACAGAGGTTGGCCAAACCGGAGGGGAGCTAATCTCAGAAACGC 1259  
DB 1121 CTACAAATGGGTGGAGACAGTGGGCAAGCAAGCATGTCAGACTAATCTCCAAAGC-C 1179  
QY 1260 GTCTGTGTCCGATCGGAGTCTGCAACTGCACTCCGTGAAGTGGGAATCGCTAGTATCG 1319  
DB 1180 ATCTCAGTTCGAGTTCGACTCTGCAACTCGAGTGAATGAAGTTGGAATCGCTAGTATCG 1239  
QY 1320 CGGATCAGCATGTCCGGGTGAATACGTTCCGGGCTTGTGACACACGCGCCGTACACCA 1379  
DB 1240 CAGATCAGCATGTCCGGGTGAATACGTTCCGGGCTTGTGACACACGCGCCGTACACCA 1299  
QY 1380 TGGAGATGGGTTTCAACAGAGGATGATCTTAACGTTAAGAGAGGCGCTTGCACGGTG 1439  
DB 1300 TGGAGATGGGTTTCAACAGAGGATGATGCGCTTAACGTTAAGAGAGGCGCTTGCACGGTG 1359  
QY 1440 AGATTATGACTGGGGTG 1457  
DB 1360 GGGTCAGCGACTGGGGTG 1377

RESULT 53  
US-10-954-147-15  
/ Sequence 15, Application US/10954147  
/ Publication No. US20050289672A1  
/ GENERAL INFORMATION:  
/ APPLICANT: CAMBIA  
/ TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells  
/ FILE REFERENCE: CAMBIA 414  
/ CURRENT APPLICATION NUMBER: US/10/954,147  
/ CURRENT FILING DATE: 2004-09-28  
/ PRIOR APPLICATION NUMBER: 60/583,426  
/ PRIOR FILING DATE: 2004-06-28  
/ NUMBER OF SEQ ID NOS: 83  
/ SOFTWARE: PatentIn version 3.2  
/ SEQ ID NO 15  
/ LENGTH: 1433  
/ TYPE: DNA  
/ ORGANISM: A. tumefaciens EHAI05  
US-10-954-147-15

Query Match 57.1%; Score 832; DB 10; Length 1433;  
Best Local Similarity 77.8%; Pred. No. 4e-248;  
Matches 1081; Conservative 0; Mismatches 265; Indels 46; Gaps 5;

QY 246 CGACGATCAGTAGTGTGCTGAGAGACGACGACCACTGGGACTGAGACACGCGCCA 305  
DB 186 CGACGATCAGTAGTGTGCTGAGAGAGATGATCAGCACTTGGGACTGAGACACGCGCCA 245  
QY 306 GACTCTTACGGGAGGACAGAGTGGGGAATTTTGGCAATGGGCGCAAGCTGATCCAGCA 365  
DB 246 AACTCTTACGGGAGGACAGAGTGGGGAATTTTGGCAATGGGCGCAAGCTGATCCAGCC 305  
QY 366 ATGCGCGTGAATGAAGAAAGGCTTCCGGTTTGAAGCTCTTTCAGTCGAGAAAGAAAG 425  
DB 306 ATGCGCGTGAATGAAGAGGCTTTCAGTTTGAAGCTCTTTCAGTCGAGAAAGTA --- 362  
QY 426 TTACGGTAATTAATCGTGAACCCATGACGCTATCGACAGAAAGCAACGGCTTAACCTAGT 485  
DB 363 -----ATGACGGTAATCGGAGAAAGAACCCCGGCTTAACCTAGT 400  
QY 486 GCCACAGCCCGCGGTAAATCTGAGGCTGCAAGCTTAAATCGGAATTAATCTGGGCGTAAAG 545  
DB 401 GCCACAGCCCGCGGTAAATCGAAGGGGCGTACCGTGTTCGGAATTAATCTGGGCGTAAAGC 460  
QY 546 GTGCCAGCGCGGCTTGTAAAGTCAAGTGAATCCCGGGCTTAAACCTGGGAATTGGCT 605  
DB 461 GCAAGTAGCGGATTTTAAGTCAAGGCTGAATCCCGAGCTCAACTGCGGAATCGCT 520  
QY 606 TTGAATCTAACAAGCTAGAGTGTGCAAGAGGAGTGAATTCATGTGTAGCACTGAAT 665  
DB 521 TTGATATCGGGATCTTGAATGGAAGAGTGAATGGAATTCGAGTGTAGAGGTGAAT 580  
QY 666 TCCGTAGAGATATGGAAGAACATCGATGCGGAAGGACCTCTCTGGTTAACTGACGCG 725  
DB 581 TTCTGATGATATTCGAGAGAACACAGATGCGAGGCGGCTTACTGTCCATTAATGACGCG 640  
QY 726 TCATGACAGAAACGTTGGGAGCAACAGATTAATACCCGTGATGTCACGCGCTTAA 785  
DB 641 TGAAGTGCAGAAACGTTGGGAGCAACAGATTAATACCCGTGATGTCACGCGCTTAA 700  
QY 786 CGATGTCACATGATGTTGGGCTTATTAAGCTTGTGAACGAAGCTTAACGCTGAAGTTG 845  
DB 701 CGATGAATGTTAGCGGTGGGCAATTAATCTGTTCGGTGGCGCAGCTTAACGCTTAACAT 760  
QY 846 ACCGCTTGGGAGTACGCTGCGAAGATTAACTCAAGGAATGACGCGGACCCGACCA 905  
DB 761 TCCGCTTGGGAGTACGCTGCGAAGATTAACTCAAGGAATGACGCGGACCCGACCA 820  
QY 906 ACGGTGATTAATGAGATTAATGAGTGAACAGGAAACCTTAACCTACCTTGAACAT 965  
DB 821 ACGGTGAGCATGTGTTTAATGAGCAACGCGCAACCTTAACCTACCTTGAACAT 880  
QY 966 -----GTACGAATTTTCTAGAGATGATGATGCTTCGGGAACGCTAACAGAGTCTG 1020  
DB 881 TCGGGGTAATGGGATTTGGAGACATGTCCTTAGTAAAGCTGGCCCCAAGAACAGGTGCTG 940  
QY 1021 CATGGCTGTCAGCTCGTGTCTGATGATGTTGGGTTAAAGTCCGCAACAGCGCAACC 1080  
DB 941 CATGGCTGTCAGCTCGTGTCTGATGATGTTGGGTTAAAGTCCGCAACAGCGCAACC 1000  
QY 1081 CTTGTCAATTAATGGCATCATTTGGTGGGCACTTTAATGAGACTCGCGGTGCAAAACG 1140  
DB 1001 CTCGCCCTTAGTGGCCAGCATTTAGTTGGGCACTTAAGGGGACTGCGCGGTGTAAGCGG 1060  
QY 1141 -GAGGAAGGTGGGATGACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACGTA 1199  
DB 1061 AAGGAAGGTGGGATGACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACACGTA 1120  
QY 1200 ATACAATGGCGCGTACAGAGGTTGGCCAAACCGGAGGGGAGCTAATCTCAGAAACGC 1259  
DB 1121 CTACAAATGGGTGGAGACAGTGGGCAAGCAAGCATGTCAGACTAATCTCCAAAGC-C 1179  
QY 1260 GTCTGTGTCCGATCGGAGTCTGCAACTGCACTCCGTGAAGTGGGAATCGCTAGTATCG 1319  
DB 1180 ATCTCAGTTCGAGTTCGACTCTGCAACTCGAGTGAATGAAGTTGGAATCGCTAGTATCG 1239  
QY 1320 CGGATCAGCATGTCCGGGTGAATACGTTCCGGGCTTGTGACACACGCGCCGTACACCA 1379



TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells  
FILE REFERENCE: CAMBIA 414B  
CURRENT APPLICATION NUMBER: US/10/953,392  
CURRENT FILING DATE: 2004-09-28  
PRIOR APPLICATION NUMBER: US 60/583,426  
PRIOR FILING DATE: 2004-06-28  
NUMBER OF SEQ ID NOS: 83  
SOFTWARE: Patentin version 3.2  
SEQ ID NO 10  
LENGTH: 1368  
TYPE: DNA  
ORGANISM: P. myrsinacearum Cambia isolate  
US-10-953-392-10

Query Match 57.0%; Score 830.8; DB 10; Length 1368;  
Best Local Similarity 78.7%; Pred. No. 9,4e-248;  
Matches 1085; Conservative 0; Mismatches 247; Indels 46; Gaps 6;

QY 66 GCGAGTGGCGGACGGGTGATGATCGGACGATCCAGAAAGGGGGTAAACGA 125  
DB 31 GGGGAGCGGACAGCGGGTGAATGACCGTGGGAATTCATCTCTACGGAATTAACGA 90  
QY 126 TCGAAAGATGTGCTAATACCGCATATCTTAAGAGGAAAGCAGGGGATCGAAAGACT 185  
DB 91 TGGAAACGTGTCTAATACCGTATACGTCTCGGGA-----GAAAGATT 136  
QY 186 TCGCGCTTTGAGCGCGCATGTCTGATTAGCTAGTGGTGGGTAAAGCCTTACCAAG 245  
DB 137 ATCGAGATGATGAGCCCGCGTGTGATTAGCTAGTGGTGGGTAAAGCCTTACCAAG 196  
QY 246 CGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 305  
DB 137 GACCTCTACGGGAGGACGAGCATGTGGGGAATTTGACAAATGGGCGCAAGCCTGATCCAGC 316  
QY 306 GACTCTACGGGAGGACGAGCATGTGGGGAATTTGACAAATGGGCGCAAGCCTGATCCAGC 365  
DB 257 GACTCTACGGGAGGACGAGCATGTGGGGAATTTGACAAATGGGCGCAAGCCTGATCCAGC 316  
QY 366 ATGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 425  
DB 317 ATGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 373  
QY 426 TTACGTAATATATCGTGAACCGGTATCGACAGAGAGACCGGGCTAACTACCT 485  
DB 374 -----ATGACGCTAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 411  
QY 486 GCGACGACCGCGGTATATCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545  
DB 412 GCGACGACCGCGGTATATCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 471  
QY 546 GTGCGCAGGCGCGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 605  
DB 472 GCGACGACCGCGGTATATCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531  
QY 606 TTGAAACTACGAGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 665  
DB 532 TTGATATCTGTGTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591  
QY 666 TCGGTAGATATGAGAGACATCGATGCGAGGAGCGCTCTGCGGTAACTGACGCG 725  
DB 592 TTCTGATATATTCGAGAGACACAGTGGCGAGAGGCGGCTCACTGCGCTGATATGACGCG 651  
QY 726 TCGATGACGAAAGCGTGGGAGCAACAGAGATTAATACCTGCTGATGATGATGATGATGATGATGATGAT 785  
DB 652 TGAGGTGCGAAAGCGTGGGAGCAACAGAGATTAATACCTGCTGATGATGATGATGATGATGATGATGAT 711  
QY 786 CGATGCACTAGTGTGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845  
DB 712 CTATAGAGACTAGCGGTGCGGAGATTAATCTGTTGCGTGGCGCAACAGGCTTAAGCTC 771  
QY 846 ACCGCTGGGAGATACGCTGCGAAGATTAACTCAAGAGAAATGACGCGGACACCGGACCA 905  
DB 772 TCGCGCTGGGAGATACGCTGCGAAGATTAACTCAAGAGAAATGACGCGGAGCGGACCA 831

QY 906 AGCGGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 965  
DB 832 AGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891  
QY 966 --GTAGCGAATTTTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
DB 892 CCCGATGCGGGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951  
QY 1021 CATGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
DB 952 CATGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011  
QY 1081 CTGCTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
DB 1012 CTCGCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071  
QY 1141 -GAGGAGGTGGGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199  
DB 1072 AGAGGAGGTGGGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131  
QY 1200 ATCAATGCGCGCTACAGAGGTTGCCAACCCCGAGGGGAGATTAATCTCAAGAAAGCGC 1259  
DB 1132 CTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1190  
QY 1260 GTGCTAGTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319  
DB 1191 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1250  
QY 1320 CGGAT 1379  
DB 1251 TGAT 1310  
QY 1380 TGGAGTGGGTTTCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1437  
DB 1311 TGGAGTGGGTTTCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1368

RESULT 56

US-10-954-147-10  
; Sequence 10, Application US/10954147  
; Publication No. US20050289672A1  
; GENERAL INFORMATION:  
; APPLICANT: CAMBIA  
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells  
; FILE REFERENCE: CAMBIA 414  
; CURRENT APPLICATION NUMBER: US/10/954,147  
; CURRENT FILING DATE: 2004-09-28  
; PRIOR APPLICATION NUMBER: 60/583,426  
; PRIOR FILING DATE: 2004-06-28  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 10  
; LENGTH: 1368  
; TYPE: DNA  
; ORGANISM: P. myrsinacearum Cambia isolate  
US-10-954-147-10

Query Match 57.0%; Score 830.8; DB 10; Length 1368;  
Best Local Similarity 78.7%; Pred. No. 9,4e-248;  
Matches 1085; Conservative 0; Mismatches 247; Indels 46; Gaps 6;

QY 66 GCGAGTGGCGGACGGGTGATGATCGGACGATCCAGAAAGGGGGTAAACGA 125  
DB 31 GGGGAGCGGACAGCGGGTGAATGACCGTGGGAATTCATCTCTACGGAATTAACGA 90  
QY 126 TCGAAAGATGTGCTAATACCGCATATCTTAAGAGGAAAGCAGGGGATCGAAAGACT 185  
DB 91 TGGAAACGTGTCTAATACCGTATACGTCTCGGGA-----GAAAGATT 136  
QY 186 TCGCGCTTTGAGCGCGCATGTCTGATTAGCTAGTGGTGGGTAAAGCCTTACCAAG 245  
DB 137 ATCGAGATGATGAGCCCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 196

```
Qy 246 CGAGCATCAAGTAAGTGGTCTGAGAGACGACGACCACTGGGACTGAGACACGGCCCA 305
    |||||
Db 197 CGAGCATCAAGTAAGTGGTCTGAGAGATGATCAAGCCACTGGGACTGAGACACGGCCCA 256
    |||||
Qy 306 GACTCTTACGGGAGGAGCAGTGGGGAATTTTGGACAAATGGGGGCAAGCCTGATCCAGCA 365
    |||||
Db 257 GACTCTTACGGGAGGAGCAGTGGGGAATTTTGGACAAATGGGGGCAAGCCTGATCCAGCC 316
    |||||
Qy 366 ATGCCCGCTGAGTGAAGAAAGCCTTCGGGTTGTAAAGCTTTCACTGAGAGAAAAGG 425
    |||||
Db 317 ATGCCCGCTGAGTGAAGAAAGCCTTCGGGTTGTAAAGCTTTCACTGAGAGAAAAGG 373
    |||||
Qy 426 TTAACGGTAATAATTCGTGACCCTGACGGTATGACAGAAAGAACCCGGCTTAACACT 485
    |||||
Db 374 -----ATGACGGTAAACCCGAGAAAGAACCCCGGCTTAACACTTCGT 411
    |||||
Qy 486 GCCAGCAGCCGCGGTAATTAAGTGGGTCAGAGGGCTTAATGCGAATTAAGTGGGGTAAAG 545
    |||||
Db 412 GCCAGCAGCCGCGGTAATTAAGAGGGGGCTAGCGTGTGAGATTAACTGGGGGTAAAGC 471
    |||||
Qy 546 GTSCGAGCGCGGCTTGTAAAGTGAATGTGAAAATCCCGGGCTTAACTGGGAATTCGT 605
    |||||
Db 472 GCACTGAGCGGCACTAATTAAGTCAAGGGGTGAATCCCGGGGCTCAACCCGGAATGCTCT 531
    |||||
Qy 606 TTGAAACTACAAAGCTAAGTGTGGCAGAGGGAGTGAATTCATGTGTGACAGTGAAA 665
    |||||
Db 532 TTGATTAAGTGTGAGTCTTGAAGTTCGAGAGAGTGTGAGTGAATTCGAGTGTGAGTGA 591
    |||||
Qy 666 TGGGTGAGATTAATGGAAGAAACATGATGCGAAGAGCGCTCTCGGGTTAACTGACGCG 725
    |||||
Db 592 TTGATGATTAATGCGAGAGAAACCAAGTGGGAGAGCGCTCACTGCTGATGATGAGCGC 651
    |||||
Qy 726 TCATGCAAGAAAGGTGGGAGGAGCAAGATTAAGTGAATCCCTGATGATCCAGCGCTAAA 785
    |||||
Db 652 TGAAGTGCAGAAAGGTGGGAGGAGCAAGATTAAGTGAATCCCTGATGATCCAGCGCTAAA 711
    |||||
Qy 786 CGATGTCATAGTGTGTTGGCCCTTATTAGGCTTGTGTAAAGAGCTTAACCGGTGAAGTTG 845
    |||||
Db 712 CTATGAGAGCTAGCGCGTTCGGGCAATATACCTGTTGCGGCGAGCAAAACGATTAACCTC 771
    |||||
Qy 846 ACCGCTTGGGAGAGTACGGTGGCAAGATTAAATCTCAAAAGAAATGACGGGGACCCGAC 905
    |||||
Db 772 TCGGCTTGGGAGAGTACGGTGGCAAGATTAAATCTCAAAAGAAATGACGGGGACCCGAC 831
    |||||
Qy 906 AGCGGTGAGATTATGATGATTAAATTCGATGCAAGCGGAAAAACCTTACCTAACCTTGA 965
    |||||
Db 832 AGCGGTGAGATGATGATGATTAAATTCGATGCAAGCGGAAAAACCTTACCTAACCTTGA 891
    |||||
Qy 966 --GTAGGAAATTTCTAGAGATAGATTAGT---GCTTCGGGAAACGCTTAACAGAGTCTG 1020
    |||||
Db 892 CCGATGCGCGTTACAGAGATGATGTTCTTCAATTAGCTGAGATCGGTGACAGATGCTG 951
    |||||
Qy 1021 CATGCTGTGCTGAGCTGCTGCTGATGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAAC 1080
    |||||
Db 952 CATGCTGTGCTGAGCTGCTGCTGATGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAAC 1011
    |||||
Qy 1081 CTTGTCAATTAATTTGCCATCAATTTGGTGGGACCTTTAATGAGATGCGCGGTGACAAACCG 1140
    |||||
Db 1012 CTCGCCCTTAATTTGCCATCAATTTAGTGGGACCTTTAAGGGGACCTGCGGTGATTAAGCG 1071
    |||||
Qy 1141 -GAGAGAGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGAGGTAGGGCTTCAACAGTA 1199
    |||||
Db 1072 AGAGAGAGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGAGGTAGGGCTTCAACAGTA 1131
    |||||
Qy 1200 ATACAAATGCGCGTACAGAGGGTTGCCAACCCGCGAGGGGAGAGCTAATCTCAAGAAAGCG 1259
    |||||
Db 1132 CTACAAATGAGTGTGACAGTGGGACGAGAGCCGCGAGGTGAGCTAATCTCAAGAAAGC-C 1190
    |||||
Qy 1260 GTGCTAGTCCGGAATGAGTCTGCAACTGCACTCCGTGAAAGTCCGGAATCCGCTAATGATCG 1319
    |||||
Db 1191 ATCTCAGTTCGGAATGCACTCTGCAACTCGAGTGAATGAAGTTGGAATCGCTAATGATCG 1250
    |||||
```

```
Qy 1320 CGGATCAGCATGTCGCGGTGAATTAACGTTCCCGGGTCTTGTACACACCGCCCTCACACCA 1379
    |||||
Db 1251 TGGATTAAGAAATGCAACGGTGAATTAACGTTCCCGGGCTTGTACACACCGCCCTCACACCA 1310
    |||||
Qy 1380 TGGAGTGGGTTTACACAGAGAGTGAATCTTAACCTTAAGAGAGGCGCTTGGCACGG 1437
    |||||
Db 1311 TGGAGTGGGTTTACACAGAGAGTGAATCTTAACCTTAAGAGAGGCGCTTGGCACGG 1368
    |||||

RESULT 57
US-11-273-617-9
; Sequence 9, Application US/11273617
; Publication No. US20060067924A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Margie D.
; APPLICANT: Hofacre, Charles
; APPLICANT: Harmon, Barry
; TITLE OF INVENTION: Probiotic Bacteria and Methods
; FILE REFERENCE: 18-03
; CURRENT APPLICATION NUMBER: US/11/273,617
; CURRENT FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/470,807
; PRIOR FILING DATE: 2003-05-14
; PRIOR APPLICATION NUMBER: WO us2004/015378
; PRIOR FILING DATE: 2005-04-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Lactobacillus aviaris
US-11-273-617-9
```

Query Match 57.0%; Score 830.8; DB 18; Length 1510;  
Best Local Similarity 76.5%; Pred. No. 9.6e-248;  
Matches 1073; Conservative 0; Mismatches 317; Indels 12; Gaps 4;

```
Qy 68 CGAGTGGCGGACGGGTGACGTAATGCAATCGG-AAACGTATCCAGAAAGGGGGGTAAAGCAT 126
    |||||
Db 105 CGAGTGGCGGACGGGTGACGTAATGCAATCGGTTAACTGCCCAAAAGAGGGGATTAACATTT 164
    |||||
Qy 127 CGAAAGATGTCATATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAAGAC--- 183
    |||||
Db 165 GGAACCAATGCTTAATCCGTATTAACCATATACCGCATGTCATTAATTAAGAGTGG 224
    |||||
Qy 184 -----CTTGGCTTTTGGAGCGGCGCATGTCATTAAGTGAATGTTGGGTAAAGGCTT 238
    |||||
Db 225 TTTTGTCTATGCTTTTGGATGAGACCGGCGGTATTAATTAAGTGGTAAAGGCTT 284
    |||||
Qy 239 ACCAAGGCGAGTCAAGTGAATGTTGCTGAGAGAGACGACCGCACTGGGACTGAGACA 298
    |||||
Db 285 ACCAAGGTGATGATACGTAACCGAGTTGAGAGAGCTGATTCGGCCACATGGAACCTGAGACA 344
    |||||
Qy 299 CGGCGCGAATCTCTACCGGAGGAGCAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTGTA 358
    |||||
Db 345 CGGCGCGAATCTCTACCGGAGGAGCAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTGTA 404
    |||||
Qy 359 TCCAGCAATCCCGGTGAGTGAAGAAAGCCTTCGGGTTGTAAAGCTTTCAAGTGAAGAA 418
    |||||
Db 405 TGAAGCAACCGCGGTGAATGAAGAGAGTCTTGGAGTCGTAATTTCTGTTTGAAGAA 464
    |||||
Qy 419 GAAAAGGTTACGGTAAATTAATCGTACCCATGACGGTATGACAGAAAGAGACCGGCTA 478
    |||||
Db 465 GAATATAGTAATTAATTAATGATTAATCACTGACGGTATCTTAACACCAAGTCAACGGCTA 524
    |||||
Qy 479 ACTATCGTCCAGAGCCGCGGTAAATAGTGAAGGTGCAAGGGTTAATCGAATTAAGTGGGC 538
    |||||
Db 525 ACTACGTCAGAGCCGCGGTAAATAGTGAAGGTGCAAGGGTTTCCGGAATTAATTTGGGC 584
    |||||
Qy 539 GTAAAGGAGTGGCGGCGCTTGTAAAGTGAATGTAAGTAAATCCCGGGCTTAACTTGGAGA 598
    |||||
Db 585 GTAAAGGAGTGGCGGCGCTTGTAAAGTGAATGTAAGTAAATCCCGGGCTTAACTTGGAGA 644
    |||||
```



```
QY 539 ATTGCGTTGAAACTACAAAGCTAGAGTGTGCGACAGGAGGTGGAATTCATGTGTAC 658
DB 645 TGTGCAATGGAAGCTGGAAGACTTGTGAGTGCAGAGAGGAGAGTCAATCTCATGTGTAC 704
QY 659 ACTGAAATCCGTAGAGATATGGAAGAACTGATGCGAAGGAGGCTCTGGGTTAA 718
DB 705 GGTGAATGCGTAGATATATGGAAGAACACACTGCGGAAAGCGCTCTGTGTCTGTAA 764
QY 719 CTGACGCTATGACGAAAGCGTGGGAGCAACAGATTATGATACCTGTATGTCACG 778
DB 765 CTGACGCTATGAGTTGAAAGCAATGGGTAGCGAAACAGATTATGATACCTGTATGTCACG 824
QY 779 CCTTAAAGATGTCAACTAGTGTGGGCTTATTTAGGCTTG--GTAAAGAGCTAACG 836
DB 825 CCGTAAACGATGAACTCTAGTGTGTGAGAGGTTCCGCCCTTCAGTGCACACTAACG 884
QY 837 GTGAAGTTGACCGCCCTGGGAGTACGCTGCGCAAGATTAAACCTAAAGGAATTGACGGG 896
DB 885 AATTAACATTTCCGCTGGGAGTACGACCGCAAGTTGAAACTCAAGGAAATTGACGGG 944
QY 897 ACCGCAACAGCGGTGATATATGTGATTATTCATGTCACGCAAGAAACCTTACCTAC 956
DB 945 GCCCGCAACAGCGGTGAGATGTGTTTATTCGAAGCAACGCGAAGAACTTACACG 1004
QY 957 CCTTGAATGTACCGAATTTTCTAGAGA-TAGATTAGTCTTGGGAAAGCTTACACG 1015
DB 1005 TCTTGAATCTTTTGAACCACTTAAGATTAGGTTTCCCTTCGGGAGCAAAATGACAG 1064
QY 1016 TGTGCAATGCTGTGCTGACGCTGTGTCGTAAGTGTGGTTAAGTCCCGCAACGAG 1075
DB 1065 TGTGATGTGCTGTGCTGACGCTGTGTCGTAAGTGTGGTTAAGTCCCGCAACGAG 1124
QY 1076 CAACCTTGTCTTAAATTTGSCATCATTTGGTGGGCACTTAAATGAGACTGCGGTGAC 1135
DB 1125 CAACCTTGTCTTGAAGTGTGCAATTCATTCATTTGGGCACTTGGGAGAGCTCCGGTAC 1184
QY 1136 AACCGAGAGAGGTGGGAGTGCATGCTCATGCTTCATGCTTATGGGTGAGGCTTAC 1195
DB 1185 AACCGAGAGAGGTGGGAGTGCATGCTCATGCTTCATGCTTATGCTGCGGTACAC 1244
QY 1196 CGTAAATCATGCGCGCTAAGAGGTTGCGCAACCGCGAGGGGAGCTAATCTCGAAA 1255
DB 1245 CGTAAATCATGCGCGCTAAGAGGTTGCGCAACCGCGAGGGTAAAGCTAATCTTAA 1304
QY 1256 GCGCGTGTAGTCCGATCGGATCTGCACTCGTGAAGTGTGAAATGCTAGTA 1315
DB 1305 GTGTTCTCTGATTGCGATTGCAAGCTCAACTCGCTGCAATGAGTGGATGCTAGTA 1364
QY 1316 ATGCGGATCAGATGTGCGGTGAATGCTTCCGGGCTTGTACACACCGCCGTCAC 1375
DB 1365 ATGCGGATCAGATGTGCGGTGAATGCTTCCGGGCTTGTACACACCGCCGTCAC 1424
QY 1376 ACCATGGAATGGGTTTCAACAGAGAGATGCTAACGTAAGAGGGGCGCTGTCAC 1435
DB 1425 ACCATGGAATGGGTTTCAACAGAGAGATGCTAACGTAAGAGGGGCGCTGTCAC 1484
QY 1436 GGTGAGATCATGACTGGGTTG 1457
DB 1485 GGTGAGATCATGAGGGTTG 1506

RESULT 58
US-11-055-637-83
; Sequence 83, Application US/11055637
; Publication No. US20050260619A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: DUBOIS, Jason
; APPLICANT: EDGE, Tom
; APPLICANT: MASSON, Luc
; APPLICANT: TREVORS, Jack T.
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
; CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES
```

```
; FILE REFERENCE: 2139-33US
; CURRENT APPLICATION NUMBER: US/11/055,637
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,288
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Probe for DNA array
; NAME/KEY: misc_feature
; LOCATION: (1)...(1491)
; OTHER INFORMATION: n = A,T,C or G
US-11-055-637-83

Query Match 57.0%; Score 830.2; DB 17; Length 1491;
Best Local Similarity 75.5%; Pred. No. 1,5e-247;
Matches 115; Conservative 0; Mismatches 339; Indels 22; Gaps 6;

QY 4 GAACGCTGGCGGATGCTTTTACATGCAATGCAACGCGACAGCGATGCTTGCATCTG 63
DB 4 GAACGCTGGCGGATGCTTTTACATGCAATGCAACGCGGAGAACMAAGCATCTTCTT 63
QY 64 GTGGCGAGTGTG-----CGACGGGTGAGTAATGATCGG-AACTATCCAGAAAG 113
DB 64 CGGGGGTGAAGCTTGTGCAACGAGCGGGTGAATACACGTTGGCAACCTAAGTAAAGC 123
QY 114 GGGGGTGAACGATCGAAATATGCTTAAATCCGCAATATCTTAAGAGAGAAAGCGGG 173
DB 124 TGGGATTAATCCGCGAAACCGGGCTAATACCGGATGATACATTCGTCATACGAGT 183
QY 174 ATCGAAGAC-----CTTGGCTTTTGAAGCGCGCATGTCTGATTAGCTAGTTGT 225
DB 184 TTGAAGAGCGGATATGCTGTACCTTACATGAGTGGCCCGCGCGCATTAAGTATGG 243
QY 226 GGGGTAAAGGCTTACCAAGCGGACGATCAGTATGTTGTTGAGAGGACGACCGCAC 285
DB 244 GAGTAAAGCTTACCAAGCGGACGATGCTAGCCGATCGAGGGTATTCGGCCACAC 303
QY 286 TGGGATGAGACACCGGCCGACCTCTTACGGAGAGGAGAGGGAATTTTGGACAAT 345
DB 304 TGGGATGAGACACCGGCCGACCTCTTACGGAGAGGAGAGGGAATTTTCCGCAAT 363
QY 346 GGGCAAGGCTGATCCAGCAATCCGCTGAGTGAAGAAAGGCTTGGGTTGTAAGCTC 405
DB 364 GACGAAGTCTGACGAGGACACCGCGCTGAGTGAAGGTTTGGATCGTAAATCTC 423
QY 406 TTTCACTGAGAAAGGTTACGGTA-AATATCTGACCCATGACGGATTCGACAG 464
DB 424 TGTGTTAGGAGAAACAAGTGCATTGCAATGTGTTGACACCTTGAACGTTACCA 483
QY 465 AGAAGCACCGGCTTAATGATGTCACACACCGCGTAAATACGTAAGGTGCAACGTTA 524
DB 484 GAAAGCCCGGCTTAATGATGTCACACACCGCGTAAATACGTAAGGTGCAACGTTGTC 543
QY 525 CGGAATTACTGGGCTTAAAGGTTGCGACAGCGGCTTGTAACTGCAATGTGAATCCCG 584
DB 544 CGGAATTACTGGGCTTAAAGGTTGCGACAGCGGCTTGTAACTGCAATGTGAATCCCG 603
QY 585 GGTAACTGGAATTTGCTTTGAAACTACAAAGCTAGAGTGTGCGAGAGGAGTGA 644
DB 604 GGTAACTGGAATTTGCTTTGAAACTGAGGAGCTTGTGTAACAAAGAGGAGTGA 663
QY 645 ATTCAATGTGTAGCAAGTAAATGCTGAGATATGGAAGAACTCATGTGCGAAAGCAGC 704
DB 664 ATTCAATGTGTAGCAAGTAAATGCTGAGATGTGGAAGAACTCATGTGCGAAAGCAGC 723
QY 705 CTCTGGGTTAAACATGAGCTCATGACGAAAGCGGGGAGCAACAGATTAGTATAC 764
```



Db	724	TCCTCGGTCTGTAACTGACCGCTGAGGTGCGCAAAACCGTGGGTAGCGAAACAGATTAGTAC	783
Qy	765	CTGTGTAGTCCACGCCCTTAAACGATGTCACTAGTTGT--GGGCTTTATTAGCTTGT	822
Db	784	CTGTGTGTTCACGCCCGTAAACGATGATGTCTAGTGTATGGGGGTTTCCGCCCTTAGT	843
Qy	823	AACGAGCTAACCGGTGAATTGACCCGCTTGGGGAGTACCGTGGCCAGATTAAACCTCA	882
Db	844	GCTTAAGTTAACGATTTAAGCACTCCGCTGGGGAGTACGGTCCCAAGACTGAATCTCA	903
Qy	883	AGGAATTGAGGGGACCCGCAAGCGGTGATTTATGTGATTTAATTCATGCAACCGCA	942
Db	904	AGGAATTGACGGGAACCCGCAAGCGGTGATTTATGTGATTTAATTCATGCAACCGCA	963
Qy	943	AAACCTTACCTTACCTTGAATGTAGCGAATTTCTAGAGATGATTAAGT-GCTTCGGG	1001
Db	964	AGAACCTTACCAAGTCTTGAATCTCTGACGCCCTTAGAATAGAGGAGTTCCCTTCGGG	1023
Qy	1002	AACGCTAACACAGGTCTGCAATGGCTGTCTGACGTCTGTCTGTGAGATGTTGGGTTAAG	1061
Db	1024	GACAGAGTGAAGAGGTGTGCAATGGTGTCTGACGTCTGTCTGTGAGATGTTGGGTTAAG	1083
Qy	1062	TCCCGCAACGAGCGCAACCTTGTCTAATTTGCGATTTGGTATTTGGTGGGCACTTAAATGA	1121
Db	1084	TCCCGCAACGAGCGCAACCTTGTCTAATTTGCGATTTGGTATTTGGTGGGCACTTAAAGT	1143
Qy	1122	GACTGCGCGGTACAAACCGGAGGAGTGGGGATGACCTCAAGTCTTCATGAGCCCTTAATG	1181
Db	1144	GACTGCGCGGTACAAACCGGAGGAGTGGGGATGACCTCAATATCATGATCCCTTAATG	1203
Qy	1182	GGTAGGGCTTCAACAGTAAATCAATGCGCGCTTACAGAGGTTTCCAAACCCGCAAGGGGA	1241
Db	1204	ACCTGGGCTTACACAGTGTCTCAATAGATGGAACAAAGGGGACGCAACCCGCAAGGCCAA	1263
Qy	1242	GCTATCTCAAGAAACCGGCTGTATCCGGAATCGGATCTGCACTCCGTAAGT	1301
Db	1264	GCAATATCCATTAACCAATTTCTCAATGTGGATGCAAGCTCCGTAATGCACTCCGTAAGT	1323
Qy	1302	CGAATCTCTAGTATTCGCGATCAGATGTTCGCGGTAAATACGTTCCCGGCTTGTATC	1361
Db	1324	CGAATCTCTAGTATTCGCGATCAGATGTTCGCGGTAAATACGTTCCCGGCTTGTATC	1383
Qy	1362	ACACCGCCGCTCACACCAATGGGAGTGGGTTTCAACAAAGCAGGTATGTTAACCCTAAGG	1421
Db	1384	ACACCGCCGCTCACACCAAGAGTGGGTAACACCGAAGTCCGTGAGGTAACTTTTGG	1443
Qy	1422	AGGGCGCTTGCACAGGTGATTTCACTGAGCTGGGGTG	1457
Db	1444	AGCCAGCCGCGCAAGGTGGGACTAATGATTTGGGGTG	1479
RESULT 59			
US-10-953-392-4			
; Sequence 4, Application US/10953392			
; Publication No. US20050289667A1			
; GENERAL INFORMATION:			
; APPLICANT: CAMBIA			
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells			
; FILE REFERENCE: CAMBIA 4148			
; CURRENT APPLICATION NUMBER: US/10/953,392			
; CURRENT FILING DATE: 2004-09-28			
; PRIOR APPLICATION NUMBER: US 60/583,426			
; PRIOR FILING DATE: 2004-06-28			
; NUMBER OF SEQ ID NOS: 83			
; SOFTWARE: Patentin version 3.2			
; SEQ ID NO 4			
; LENGTH: 1395			
; TYPE: DNA			
; ORGANISM: S. meliloti 1021			
US-10-953-392-4			
Query Match			
Best Local Similarity			
56.8%; Score 827.2; DB 10; Length 1395;			
77.5%; Ptd No. 1,2e-246;			

	Matches	1084;	Conservative	0;	Mismatches	268;	Indels	46;	Gaps	5;
Qy	66	GGCGAGTGGCGGACGGGTGAGTAATGTCATCGGAAGCTATCCAGAAAGGGGGTAAACGCA	125							
Ddb	33	GGGGAGCGGACAGACGGGTGAGTAAACGGCGGGAAATCACTCTTTTCTACGGAATTAACGCA	92							
Qy	126	TGGAAGAATGTGCTAAATACCGCATATATCTTAAGAGGAAACAGGGGATTCGAACACT	185							
Ddb	93	GGGAAACTTGTGCTAAATCCGTATAGCCCTTCGGGGGAAAGATTTATCGGGAAA-----	147							
Qy	186	TGCGCTTTTGGACGGCCGATGTCTGATTAGCTAGTTGGTGGGGTTAAAGGCTTACCAAG	245							
Ddb	148	-----GATGAGCCCCCGCTTGGATTAGCTAGTTGGTGGGGTTAAAGGCTTACCAAG	198							
Qy	246	CGACGATCAGTATGTGTCTGAGAGGACGACCACTGGGACTGAGACACGGCCCA	305							
Ddb	199	CGACGATCAGTATGTGTCTGAGAGGATGATCAGCACAATTGGGACTGAGACACGGCCCA	258							
Qy	306	GACTCTTACGGAGGCGACAGTGGGGAAATTTTGGACAAATGGGGCGGAAGCTGATCCAGCA	365							
Ddb	259	AACCTCTTAAGGGAGGCGACAGTGGGGAAATTTGGACAAATGGGGCGGAAGCTGATCCAGCC	318							
Qy	366	ATGCCGCGTGAATGAAGAAAGGCGCTTGGGTTGTTAAAGCTCTTTCAGTCGAGAAAGAAAGG	425							
Ddb	319	ATGCCGCGTGAATGAAGAAAGGCGCTTGGGTTGTTAAAGCTCTTTCACCGGTAAGATA---	375							
Qy	426	TTACGGTAATATATCGTGACCCATGACCGGTATCGACAGAAAGACACCGGCTAATCTACGT	485							
Ddb	376	-----ATGACGGTAACCGGAAAGAAAGCCCGGCTAATCTTCTGT	413							
Qy	486	GCCAGACGCGCGGTAAATACGTAGGGGTGCAAGGCTTAATCGAAATTAATCTGGGGTAAAG	545							
Ddb	414	GCCAGACGCGCGGTAAATACGTAGGGGTGCAAGGCTTAAAGGCTTTTCAACCGGTAAGAC	473							
Qy	546	GTGCGGAGGCGGCTTGTAGTCAATGTGAAATGCCCGGGCTTAACTCTGGGAATTGCGT	605							
Ddb	474	GACAGTAGGCGGATTTTAAATGAGGGGGTGAATTCAGGGGCTCAACCTCTGGAATGCTCT	533							
Qy	606	TTGAAACTACAAAGCTTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTAGCAGTGA	665							
Ddb	534	TTTCAATCTGGCAATCTTAGAGTCCAGAAAGGTGAGTGAATTCGAGTGTAGAGGTGA	593							
Qy	666	TGCGTAGATATGAAAGACATGATGCGGAAGGACGCTCTGGGTTAAACATGACGC	725							
Ddb	594	TTGCTGATATTTGGAGGAACACACAGTGGGGAAGGGCGCTCACTGGTCTGAACTGACGC	653							
Qy	726	TCAATGACAAACCGTGGGAGGAAACAGGATTAGATACCCTGTAGTCAACGCTTAA	785							
Ddb	654	TGAGGTGCGAAACCGTGGGAGGAAACAGGATTAGATACCCTGTAGTCAACGCTTAA	713							
Qy	786	CGATGTCAACTAGTTTGTGGGCTTATATAGGCTTGTGTAAAGAACTPAACGCGTGAAAGTTG	845							
Ddb	714	CGATGAATGTAGCCGTGCGGCAAGTTTACGTTCGTGGGCGCAGCTPAACGATTTAAACAT	773							
Qy	846	ACCGCTTGGGAGTACGCTGCGAAGATTAAACCTCAAGGAATTGACGGGACCCGACACA	905							
Ddb	774	TCCGCTTGGGAGTACGCTGCGAAGATTAAACCTCAAGGAATTGACGGGAGCCCGACACA	833							
Qy	906	AGCGGTGGAATTATGTGATTAATTTCCATGCAAGCGGAAAAACCTTACCTAACCTTGCAT	965							
Ddb	834	AGCGGTGGAACATGTGTATTAATTTCAAGGACAGCGGCAAACTTACCAAGCTTGCAT	893							
Qy	966	GTAGCGAATTTCTAGAGA-----TAGATTAAGTCTTCGGGAACGCTTAACACAGTGTG	1020							
Ddb	894	CCCGATTCGGGATACGAGAGATGCTATCTTCAAGTTCCGGCTCGATTCGGAAGACAGTGTG	953							
Qy	1021	CATGGCTGTGCTAGCTCGTGTGCTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAAC	1080							
Ddb	954	CATGGCTGTGCTAGCTCGTGTGCTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAAC	1013							
Qy	1081	CTTGTCAATTAATGGCCATCAATTTGGTTGGGCACTTAAATAGACTGCGGGAACAAACCG	1140							
Ddb	1014	CTGCGCTTATGTCGACGATTAACGATTTGAGGCACTTAAGGGGACTGCGGCTGATTAAGCG	1073							

Qy	1141	GAGGAAGTGGGAGTGAAGTCACAACTCCATAGGCCCTTAATGGGATGGAGGCTTACACAGTA	1159
Db	1074	AGAGGAAGTGGGAGTGAAGTCACAACTCCATAGGCCCTTAATGGGATGGAGGCTTACACAGTA	1133
Qy	1200	ATACAAATGGCGGTACAGAGGGTTCCCAACCCGCGAGGGGGAGCTAACTCAGAAAGCGC	1255
Db	1134	CTACAAATGGTGGTGAACATGGGCAAGCGAGAGCCGCGAGGGTCAGCTAACTCCAAAGG-C	1192
Qy	1260	GTCTAGTCCGAGATCGAGTCTGCACTCGACTCCGTGAATCGGAAATCGCTAGTAATCG	1319
Db	1193	ATCTCAGTTCCGATTTGCACTCTGCAACTCGAGTGTAGTAAGTTGAATCGCTAGTAATCG	1252
Qy	1320	CGGATCAGCATGTCCGCGGGAATAGATTCCCGGGGCTTGTACACACCGCCGTCACACGA	1379
Db	1253	CAGATCAGATGCTCCGGTGAATACGTTCCCGGCGCTTGTACACACCGCCGTCACACGA	1312
Qy	1380	TGGAGTGGGTTTCAACGAAGACAGTATGCTTAACCTGAAGAGGCGCTTGCACGCGT	1439
Db	1313	TGGAGTTGGTTCTAACCCGAAGGTATGTGGCTTAACCGCAAGGAGGACCTAACACGCGTA	1372
Qy	1440	AGATTCAATGACTGGGCTG 1457	
Db	1373	GGGTCAAGCACTGGGCTG 1390	
RESULT 60			
US-10-954-147-4			
; Sequence 4, Application US/10954147			
; Publication No. US20050289672A1			
; GENERAL INFORMATION:			
; APPLICANT: CAMBIA			
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells			
; FILE REFERENCE NUMBER: US/10/954,147			
; CURRENT FILING DATE: 2004-09-28			
; PRIOR APPLICATION NUMBER: 60/583,426			
; PRIOR FILING DATE: 2004-06-28			
; NUMBER OF SEQ ID NOS: 83			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 4			
; LENGTH: 1395			
; TYPE: DNA			
; ORGANISM: S. meliloti 1021			
US-10-954-147-4			
Query Match 56.8%; Score 827.2; DB 10; Length 1395;			
Best Local Similarity 77.5%; Pred. No. 1.2e-246; Mismatches 268; Indels 46; Gaps 5;			
Matches 1084; Conservative 0;			
Qy	66	GGCGAGTGGCGGAGCGGTGAGTAATGCAATCGGAAGCTATCCAGAGAGGGGGGTAAACGCA	125
Db	33	GGGGAGCGGCAAGCGGTGAGTAATGCAATCGGTGGAATCTACCTTTTCTACGAAATAACGCA	92
Qy	126	TCGAAAGATGTCTAATAACCGATATACTTAAGAGAGAAAGCAGGGGATCGAAAGACCT	185
Db	93	GGGAATCTGTCTAATAACCGATAGACCTTCGGGGGGAAGATTTATCGGAGAA-----	147
Qy	186	TGCGCTTTTGAAGCGCGCGCATCTTCATAGTAATGTGGTGGGGTAAAGGCGCTACCAAG	245
Db	148	-----GGATGAGCCCGCGCTGGATTGATGTAATGTGGTGGGGTAAAGGCGCTACCAAG	198
Qy	246	CGACGATCAGTATGTTGTCGTAGAGAGACGACCGACACTGAGGACTGAGACACCGGCCA	305
Db	139	CGACGATCAGTATGTCGTGTAGAGAGATGATAGCCACTGTTGGACTGAGACACGGGCCA	258
Qy	306	GACTCTTACGAGGCGACGATGGGGATTTTGAACAATGGCGCAAGCTTATCTACGCA	365
Db	259	AACTCTTACGAGGCGACGATGGGGATTTTGAACAATGGCGCAAGCTTATCTACGCA	318
Qy	366	ATGCCCGCGTGAATGAAGAGGCTTCGGGTGTAAAGCTCTTCACTGAGAGAAAGG	425
Db	319	ATGCCCGCGTGAATGAAGGCTTCGGGTGTAAAGCTCTTCACTGAGAGAAAGT--- 375	

QY	426	TTACGGTAATAATATCGTGAACCCATGACCGGTATCGACAGAAAGAACCGCGCTAACTACGCT	485
Db	376	-----ATACGGTAAACCGAGAAAGAACCCCGCTAACTTCCGT	413
QY	486	GCCAGCACCCCGGTAATACGTAGGGGTCCAAAGCCTTAAATCCGAAATTACCTGGCGGTAAAG	545
Db	414	GCCAGCACCCCGGTAATACGAAGGGGCGTACGTTGTTCCGAATTACCTGGCGGTAAAGC	473
QY	546	GTCGCGAGGCGGCCCTTGTAACTCAGATGTGAATCCCGGGCTTAACTTGGGAATTGGCT	605
Db	474	GCAGGTAGCGGATGTGTTAAAGTGAAGGGGTGAATCCCAAGGGCTCAACCTCGGAACCTGCT	533
QY	606	TTGAAACTTACAAAGCTTAGAGTGTGCGACAGAGGAGGTGGAAATTCATGTGTACAGTGA	665
Db	534	TTCAATACCTGGCAATTTAAGTTCACAAAGAGGTGAATTCGAGTGTAAAGGTGA	593
QY	666	TGCGTAGAGATATATGGAAAGACATCGATGGCGAAAGGACGCTCTGGGTAACTACCTACGCG	725
Db	594	TTCTGTAATATTCGAGAGAACACAGTGGCGAAAGGCGCTCACTGTCGTGAATCTACCG	653
QY	726	TCATGCAAGAAAGCGTGGGAGCAAAACAGATTGAATACCTGTAGTCCACGCGCTTAA	785
Db	654	TGAGGTGGAAAGCGTGGGAGCAAAACAGATTGAATACCTGTAGTCCACGCGCTTAA	713
QY	786	CGATGTCAACTATGTTGTGGGCTTTAAGCTTGTGTAAAGCAAGCTTAACGCGTGAAGTTG	845
Db	714	CGATTAATGTTAAGCGGTGGGAGTTTACTGTTGGTGGCGAGCTTAACGCAATTAACAT	773
QY	846	ACCGCTGGGGAGTACGGTCCGCAAGTTAAACTCAAAAGAAATTGACGGGAGACCGGCACA	905
Db	774	TCCGCTTGGGAGTACGGTCCGCAAGTTAAACTCAAAAGAAATTGACGGGAGACCGGCACA	833
QY	906	AGCGGTGATTATGTGATTATTCGATGCAACGGCAAAACCTTACCTTACCTTGAAT	965
Db	834	AGCGGTGAGCATGTGTGTTAATTCGAAGCAACGGCAAAACCTTACAGCCTTGAAT	893
QY	966	GTAACGAATTTTCTAAGAC-----TAGATTAGTCTCTGGGAGCGCTAACACAGGTGCTG	102
Db	894	CCCGATCCCGATATGAGAGATCTGATCTCTTCAATTCGGCTGATTCGAGAACAGGTGCTG	953
QY	1021	CATGCTTCCGTAACGCTGTGTCGGAATGTTGGGTTAAGTCCCGCAACGAGCGCAAC	108
Db	954	CATGCTTCCGTAACGCTGTGTCGGAATGTTGGGTTAAGTCCCGCAACGAGCGCAAC	101
QY	1081	CTTGTGATTAATTTGTCATCATTTTGGTTGGGCACTTTAAATGAGACTGCGGTGACAAACCG	114
Db	1014	CTCGCCCTTAATGTTGCGACGATTCAGTTGGGCACTTAAGGGGACTGCGGTATTAAGCGG	107
QY	1141	-GAGAGAGTGGGAGTGAAGTCMAAGTCTCATGAGCCCTTATGAGTGAAGGCTTTCACAGTA	119
Db	1074	AGAGAAAGTGGGAGTGAAGTCMAAGTCTCATGAGCCCTTACGGGCTGGGCTTACACAGTGA	113
QY	1200	ATACATAGCGCGTACAGAGGTTTGCCAACCCGCGAGGGGGAGCTTAATTCAGAAAGCGC	125
Db	1134	CTACATAGTGTGTGACAGTGGCACCGCAACCGCAGAGTGAAGCTTAATTCACAAAGC-C	119
QY	1260	GTCGTAATCCCGATCGGAGTTCGCAACTCGACTCGTGAAGTCCGAAATCGCTAATGATCG	131
Db	1193	ATCTCAATTCGGAATTCGACTCTGCAACTCGAATGATGAAGTTGGAATTCGCTAATGATCG	125
QY	1320	CGGATCAGCATGTCCGCGTGAATTAAGTTCGCGGCTTGTATACACACGCGCCGTACACCA	137
Db	1253	CAGATCAGGATATCTCGCGGTAATAGTTCGCGGCTTGTATACACACGCGCCGTACACCA	131
QY	1380	TGGAGTGGGTTTCAACGAGACAGGTAGTCTTAACCTTAAGAGGGGCGTTTGCACAGGTG	143
Db	1313	TGGAGTGGGTTTCTAACCGAAGGTAGTGCCTTAACCGCAAGAGGCGCAAGCTTAACAGGTGA	137
QY	1440	AGATTCAATGACTGGGGTG 1457	
Db	1373	GGGTACGCACTGGGGTG 1390	

```
RESULT 61
US-11-055-637-74
; Sequence 74, Application US/11055637
; Publication No. US20050260619A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: DUBOIS, Jason
; APPLICANT: EDGE, Tom
; APPLICANT: MASSON, Luc
; APPLICANT: TREVORS, Jack T.
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
; FILE REFERENCE: 2139-3305
; CURRENT APPLICATION NUMBER: US/11/055,637
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,288
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 1535
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Probe for DNA array
US-11-055-637-74

Query Match      56.7%; Score 825.4; DB: 17; Length 1535;
Best Local Similarity 75.0%; Pred. No. 4,66-246;
Matches 1103; Conservative 0; Mismatches 351; Indels 17; Gaps 5;

Qy      4  GAACGCTGGCGGCGATGCTTTACATGCAAGTCGACGAGCGGACGAGTGGCTTCATCTG 63
Db      24  GAAGCGTGGCGGCGGCTGCTTAATACATGCAAGTCGACGAGCGGACGAGTGGCTTCATCTG 83
Qy      64  GTGGCGAGTGGCGGCGGAGTGAATGATCGAGTGG-ACGTAATCGAAGAGGCGGAGTAC 122
Db      84  GAAGTTAGCGGCGGAGTGAATGATCGAGTGG-ACGTAATCGAAGAGTGAATGAGTGGAGTAC 143
Qy      123  GCATCGAAGATGCTTAATACCGCAT------TACTTAAGAGAGAAACAGAGG 173
Db      144  TTCGGGAAACCGGAGCTAATACCGGATTAATCTTTCTCTCGAGAAAGGTTGAAAG 203
Qy      174  ATCGAAGACCTTGCGGCTTTTGAAGCGGCGGATGCTGATTAAGTATGTTGGGCTTAA 233
Db      204  ACGGCTTGCGGCTGCTTAATCAAGATGGGCGCGCGGCGATTAAGTATGTTGGGCTTAA 263
Qy      234  GGCTTACCAAGCGGCGGATCGATGATGTTGCTGAGAGGACGACGACCACTGGGACTG 293
Db      264  GGCTTACCAAGCGGCGGATCGATGATGTTGCTGAGAGGAGTATGGGCGACACTGGGACTG 323
Qy      294  AGACACGGCGCGGAGTCTTACGAGAGGAGGAGGAGTGGGAAATTTTGAACAATGGGCGAAG 353
Db      324  AGACACGGCGCGGAGTCTTACGAGAGGAGGAGGAGTGGGAAATTTTCCGCAATGGAGAAAG 383
Qy      354  CCTGATCCAGCAATGCGCGGAGTGAAGAGGCGCTTGCGGCTGTAAGTCTTTCACTC 413
Db      384  TCTGACGGAGCAACGCGCGGAGTGAAGAGGCTTTGGAATCGTAATCTCTGTTATCA 443
Qy      414  GAGAAAGAAAGTTACGCTTAATATCGTGAACCCATGACCGATTCAGACAGAAAGACAC 473
Db      444  GGGAAAGAAAGTATCGGAGTATACGCGGATACCTTGAACGATACGACAGAAAGCAC 503
Qy      474  GGCTTAATACGTCGAGCAGCGCGGCTTAATACGTAAGGCTCAAGCGTTAATCGAATTAC 533
Db      504  GGCTTAATACGTCGAGCAGCGCGGCTTAATACGTAAGGCTCAAGCGTTCGGAATTAT 563
Qy      534  TGGGCGTAAAGGCTGCGAGCGGCGCTTGTATGATGATGTAATCCCGGCGCTTAAC 593
Db      564  TGGGCGTAAAGGCTGCGAGCGGCGCTTGTATGATGATGTAATCCCGCGCTTAAC 623
Qy      594  TGGGAATTGCTTGAATCAAAAGCTAAGTGTGCGAGAGGAGGAGTTCATGAT 653
```

```
Db      624  GTGGAAGCTCATTTGAAACTGGGAGCTTGAAGTGAAGAGAGAGAGCGGAATTCACAGT 683
Qy      654  GTAGCATGAATATGCTTAAGATATGGAAGAACTATGATGGCGAGCGCTCCGGGT 713
Db      684  GTAGCGGTGAATATGCTTAAGATATGGAAGAACTATGATGGCGAGCGCTCTTGTGTC 743
Qy      714  TAACTGAAGCTCATGAGCAAGAAAGGCTGGGAGCAAAAGATTAATACCTGTGATG 773
Db      744  TGAATGAGCTGAGGCGCGAAGCGTGGAGCGCAACAGATTAATACCTGTGATG 803
Qy      774  CCAAGCGCTTAAGCATATGCTAATGTTGGGCTTAATAGCC--TTGGTAAAGAACT 831
Db      804  CCAAGCGCTTAAGCATATGCTAATGTTAGAGGCTTCCGCTTAATGCTGAGT 863
Qy      832  AACCGGTAAATGATACCGCTGGGAGTACGCTGCGAAGTAACTAAAGAAATGTA 891
Db      864  AACGCAATTAAGCATCTCGCTGGGAGTACGCGCGAAGGCTGAATCTAAAGAAATGTA 923
Qy      892  CGGGGACCGGCAAGCGGCTGATTAATGGAATTAATTCGATCAACGCGAAGAACTTA 951
Db      924  CGGGGACCGGCAAGCGGCTGAGTATGTTGTTAATTCGAAGCAACGCGAAGAACTTA 983
Qy      952  CTAACCTTGAATGATGAGCAATTTTCTAGAGTAG--ATTAGTCTTCGGAAGCTTA 1008
Db      984  CCAAGCTTGAATGATGAGCAATTTTCTAGAGTAG--ATTAGTCTTCGGAAGCTTA 1043
Qy      1009  ACAAGGTGCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
Db      1044  TGAAGGTGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103
Qy      1069  ACGAGCGCAACCTTGTGATTAATGTCATCATTTGGTGGGCACTTAATAGACTGCC 1128
Db      1104  ACGAGCGCAACCTTGTGATTAATGTCATCATTTGGTGGGCACTTAATAGACTGCC 1163
Qy      1129  GGTGACAAACCGGAGAGGTTGGGATGACGTCAGTCTCATATGCGCTTAATGAGTACG 1188
Db      1164  GGTGACAAACCGGAGAGGTTGGGATGACGTCAGTCTCATATGCGCTTAATGAGTACG 1223
Qy      1189  CTTCAACAGTAAATCAATGCGCGGTGACAGAGGTTGCGCAACCGCGAGGGGAGCTATC 1248
Db      1224  CTTCAACAGTAAATCAATGCGCGGTGACAGAGGTTGCGCAACCGCGAGGGGAGCTATC 1283
Qy      1249  TCGAAGACCGGCTGATGCTCGGATCGGATCGGATCGGATCGGATCGGATCGGATCG 1308
Db      1284  CCAATAAACCAATCTCAGTTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1343
Qy      1309  GCTAATATGCGGATGACGATGCTGCGGATTAATGCTCCGAGCTCTTGTACACACCG 1368
Db      1344  GCTAATATGCTGATGACGATGCTGCGGATTAATGCTCCGAGCTCTTGTACACACCG 1403
Qy      1369  CCGTCAACCATGAGATGAGTTCACCAAGAGGATGATCTAACCGTAA--GAGAGGC 1426
Db      1404  CCGTCAACCATGAGATGAGTTCACCAAGAGGATGATCTAACCGTAA--GAGAGGC 1463
Qy      1427  GCTTGCACCGGTGATTCATGACTGGGCTG 1457
Db      1464  GCCGCCGAGGTGGAACGATGATGGGCTG 1494

RESULT 62
US-11-084-508-3
; Sequence 3, Application US/11084508
; Publication No. US20050260737A1
; GENERAL INFORMATION:
; APPLICANT: Rahman, Raja Noor Zaiha Abd.
; APPLICANT: Salleh, Abu Bakar
; APPLICANT: Basri, Mahlihan
; APPLICANT: Hun, Chin John
; TITLE OF INVENTION: Novel Lipase Gene from Bacillus sphaericus 205y
; FILE REFERENCE: KAN-101
; CURRENT APPLICATION NUMBER: US/11/084,508
; CURRENT FILING DATE: 2005-03-18
```

```

? PRIOR APPLICATION NUMBER: MY 2004-0958
?
? PRIOR FILING DATE: 2004-03-18
?
? NUMBER OF SEQ ID NOS: 27
?
? SOFTWARE: PatentIn version 3.2
?
? SEQ ID NO 3
?
? LENGTH: 1459
?
? TYPE: DNA
?
? ORGANISM: Bacillus sphaericus
?
? US-11-084-508-3

```

Query Match	56.6%	Score 825	DB 17	Length 1459
Best Local Similarity	76.5%	Pred. No. 6.1e-246		
Matches 1094	Conservative 0	Mismatches 320	Indels 17	Gaps 6

Qy	4	GAAGCGTGGCGGATGCTTTACATGACGAAGTCGAAC - GCAGCAGCGATCTTGACATC -	61
Db	29	GAAGCGTGGCGGCTCCCTTAATACATGCAAGTGCAGCAACGAAGAAAGAACCTTGCTCCT	88
Qy	62	TGTGCGAGTGGCGGACGGGTGATGATATCGG - AAAGTATCAGAAAGAGGGGGTGA	120
Db	89	TTGACGTTTACCGCGGCGACGGGTGATGATACAGTGGGCAACTTACCTTTATGTTTGGGATA	148
Qy	121	ACCGATGGAAGATATGTCCTAATACCGCATATCTCTAAGACGAAGAAAGCAGGGGATTCGAA	180
Db	149	ACTCGGGAAACCGGGGCTAATACCAATTAATCTATTTCATTCAATGATGTAATAATCTGAA	208
Qy	181	GAC-----CTTGGCTTTTGGACGCGCGGATGTCATTTACTAGTGGTGGGGTGA	231
Db	209	AGACGATCTCGGCTGTCTGCTAATGAATGGGCCCGGGCGCATTTAGCTGATTTGGTGAAGTGA	268
Qy	232	AAGGCTTACCAAGGCCATCATGATGTTGCTTGAAGAGACACCACTATGGGAC	291
Db	269	ACGGCTTACCAAGGCCACATGATGTCGACCTGAGAGGGATGATCGGCCACTATGGGAC	328
Qy	292	TGAGACACGGCCAGACTCTCTACGGGAGGACACAGTGGGGAAATTTTGACATAGGCGCA	351
Db	329	TGAGACACGGCCAGACTCTCTACGGGAGGACACATGAGGAATCTTCCACATATGGCGCA	388
Qy	352	AGCCTGATCAGCAATGCGCGTGAATGAAGAAGCCTTCGGGTTGTAAAGCTCTTTGAC	411
Db	389	AGCCTGATGAGAGCAAGCGCGTGAATGAAGAAGGTTTCCGATGTAAACTCTGTGTTG	448
Qy	412	TCGAGAAAGAAAGGTTAACGGTAAATTAATGTAACCATAGACGGTATGCACAGAAAGCA	471
Db	449	AAGGAAAGAAAGGTTAACGGTAAATTAATGTAACCATAGACGGTATGCACAGAAAGCC	508
Qy	472	CCGGCTACTACGTGCGCAGACGCCCGGATATACGTAAGGTGCAAGCGTTAATCGAATT	531
Db	509	ACGGCTACTACGTGCGCAGACGCCCGGATATACGTAAGGTGCAAGCGTTTCCGAATT	568
Qy	532	ACTGGGCTTAAAGGGTGGCGACGGCGGCTTGTAAATGTCAGATGTGAATCCCGGGCTTAA	591
Db	569	ATTGGGGGTAAACCGCGCGACGGCGGATCTTTAAGTCTGAATGTGAATCCCAAGGCTTAA	628
Qy	592	CCTGGGAATTGCGTTTGAACCTACAAAGCTTAAGTGTGGCAGAGGGAAGTGAATTCAT	651
Db	629	CCGTGAGAGGATCATTTGAAACTGCGGGGAACTTGAATGTGCAAGAGGAAGTGAATTCGA	688
Qy	652	GTTGAGCATGAATATGCTGATGAGATATGGAAGAACATGATGGCGAAGGACGCTCTGG	711
Db	689	GTTGAGCATGAATATGCTGATGAGATATGGAAGAACACATGATGGCGAAGGACCTTCTGG	748
Qy	712	GTTTAAACCTGACGCTCATGACGAAAGCGTGGGGAGCAACAGATTAGATACCTGTGA	771
Db	749	TCTGTAACTGACGCTCATGACGCGAAGACGTTGGGAGCAACAGATTAGATACCTGTGA	808
Qy	772	GTCACAGCCCTTAAACGATGTCACTAGTTGT - -GGGCTTATTTAGCTTGTGAACCAAG	829
Db	809	GTCACAGCCGTTAAACGATGTGCTTAAAGTTTAAAGGGGTTTCCGCTTATGTCGTGAC	868
Qy	830	CTTAAACGCTGAATGTGACCGGCTGGGGAGTACGCTGCGCAAGTTTAAACTTAAAGAAATT	889
Db	869	CTTAAACGATTTAACCACTCGGCTGGGGAGTACGCTGCGCAAGCTTAAACTTAAAGAAATT	928

OY	890	GACGGGACACCCGACAAACGGGGGATTTATGAGATTAATTCGATGCAACGGCAAAAACCT	969
Db	929	GACGGGGGGCCCGACAAAGCGGTGGAGCATGGGTTTATTTTCGAAGCAACCGAAGACCT	988
OY	950	TACCTTACCCCTTGAACAT--GTAGCCAAATTTTCTAGAGATAGATTAGTCTTCGGGAACCG	1000
Db	989	TATCAAGTCTTGACATCCCGTTGACCACTGTATAGATATATGGTTTTCCTTCGGGGACAAC	104
OY	1007	TAAACAAGGTGCTGCATGAGCTGTGCTCAAGCTCTGTGTGTGTGATAGATGTTGGGTTAAATCCG	106
Db	1049	GGTGACAGGTGGTGATGGTTGTCTGTACGCTCTGTGTGTGTGATAGATGTTGGGTTAAATCCG	110
OY	1067	CAACGAGCCCAACCCCTTGACATTAATGACATCAATTTGGTGGGCACTTTATATAGACTG	112
Db	1109	CAACGAGCCCAACCCCTTGATCTTATGTTGCATCAATTTATGTTGGGCACTTAAAGTACTG	116
OY	1127	CCGGTACAAACCCGAGGAGAGGTTGGGGATGACGTCAAGTCTCTCATGACCCTTATGGGATG	118
Db	1169	CCGGTACAAACCCGAGGAGAGGTTGGGGATGACGTCAATATCATATGCCCCCTTATGACCTG	122
OY	1187	GGCTTACACGTTAATPACAATGGCCGCTPACAAGGTTGCCAACCCCGGAGGGGAGCTAA	124
Db	1229	GGTTACACACGCTGCTACATGACATGACATCAAAACGTTGCCAACTCGGAGAGGAGACTAA	128
OY	1247	TCCTCAAAAGCCGGCTGTAATGTCGGGATGCGAGTCTGCAACTCGACTCCGTAAGTGGAA	130
Db	1289	TCCGATTAAGTGTGTTCTCAGTTCTGGATTTAGCTGCAACTCGCTTACATGAAGCCGGAA	134
OY	1307	TCGCTAGTAATCGCGGATCAGCATGTCCGGTGAATACGTTCCCGGCTCTTGATACACC	136
Db	1349	TGCTGTAGTAATCGCGGATCAGCATGTCCGGTGAATACGTTCCCGGCTCTTGATACACC	140
OY	1367	GGCCGTCACACCATGGGAGTGGGTTTCCACGAAGCAGTATGCTTAACCT	1417
Db	1409	GGCCGTCACACCATGAGGTTTGTAAACACCGGAAGTGGTGGGTAACCTT	1459
RESULT 63			
US-11-045--004-1			
Sequence 1, Application US/11045004			
Publication No. US20060078901A1			
GENERAL INFORMATION:			
APPLICANT: BUCHRIESEN, CARMEN			
APPLICANT: FRANGEUL, LIONEL			
APPLICANT: COUVE, ELISABETH			
APPLICANT: RUSNIOX, CHRISTOPHE			
APPLICANT: FSIHI, HAFIDA			
APPLICANT: DEHOUX, PIERRE			
APPLICANT: DUSSURGET, OLIVIER			
APPLICANT: CHEYOUNI, FARID			
APPLICANT: MEDJARI, HAFED			
APPLICANT: GLASER, PHILIPPE			
APPLICANT: KUNST, FRANCK			
APPLICANT: COSSART, PASCALE			
APPLICANT: DANIELS, JUSTIN			
APPLICANT: GOEBEL, WERNER			
APPLICANT: KREFT, JURGEN			
APPLICANT: KUHN, MICHAEL			
APPLICANT: NG, EVA			
APPLICANT: VAZQUEZ-BOLAND, ANTONIO			
APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO			
APPLICANT: GARRIDO-GARCIA, PATRICIA			
APPLICANT: TIERREZ-MARTINEZ, ALBERTO			
APPLICANT: AMEND, ALEXANDRA			
APPLICANT: CHAKRABORTY, TRINAD			
APPLICANT: DOMANN, EUGEN			
APPLICANT: HAIN, THORSTEN			
APPLICANT: BERGEH, PATRICK			
APPLICANT: CHARBIT, ALAIN			
APPLICANT: DURANT, LIONEL			
APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO			
APPLICANT: BAQUERO, FERNANDO			

```
APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
APPLICANT: GOMEZ-LOPEZ, NORIA
APPLICANT: MADUENIO, ENCARNIA
APPLICANT: PABLOS, BETRIZ DE
APPLICANT: WEHLAND, JURGEN
APPLICANT: KARST, UWE
APPLICANT: ENTIAN, KARL-DIETER
APPLICANT: HAUF, JORG
APPLICANT: ROSE, MATTHIAS
APPLICANT: VOSS, HAMUT
TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
FILE REFERENCE: 05394.0018-02
CURRENT APPLICATION NUMBER: US/11/045.004
PRIOR FILING DATE: 2005-01-28
PRIOR APPLICATION NUMBER: 10/637,657
PRIOR FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: 10/257,023
PRIOR FILING DATE: 2002-10-08
PRIOR APPLICATION NUMBER: PCT/FR01/01118
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: FR 00/04,629
PRIOR FILING DATE: 2000-04-11
NUMBER OF SEQ ID NOS: 2854
SOFTWARE: PatentIn version 3.3
SEQ ID NO 1
LENGTH: 2944528
TYPE: DNA
ORGANISM: Listeria monocytogenes
US-11-045-004-1

Query Match      56.6%; Score 824.8; DB 18; Length 2944528;
Best Local Similarity 76.3%; Pred. No. 4.3e-245;
Matches 1121; Conservative 0; Mismatches 332; Indels 17; Gaps 8;

QY      4      GAAAGCTGGGGGCGCTGTTTACATGCAAGTGGAAAGGCGAGCGAGTGGCTTGCATCT- 62
DB      237499  GAAAGCTGGGGGCGCTGTTTACATGCAAGTGGAAAGGCGAGCGAGTGGCTTCTT 237558

QY      63      -GGTGGGAGTGGGCGGAGCGGGTGAATGATCGG-AAAGTATCCGAAGAGGGGGTA 120
DB      237559  CCAATGTTAGTGGGAGAGGGGTGAGTAAACAGTGGGCACTGCTGTAAGTTGGGATA 237618

QY      121     ACGCATGAAAGATGTGCTATACCGCAT-ATACTTAAAGAGAAAGCAGGGATCGAA 179
DB      237619  ACTCGGGGAAACCGGGGCTAATACCGATGATTAAGATGTGGCGCATCGCCTTGAA 237678

QY      180     AGAC-----CTTGGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTA 231
DB      237679  AGATGTTTTCGGCTATGCTTACAGATGGGCGCGGTCATTTAGCTAGTTGGTGGGTA 237738

QY      232     AAGGCTTACCAAGGCGAGCATCAGTATGTTGTTGAGAGAGCAGCAGCCACACTGGAC 291
DB      237739  ATGCTTACCAAGGCGAGCATGCAATGCGACCTGAGAGGGTGAATCGGCCACTGGAC 237798

QY      292     TGAGACACGGGCCAGACTCTTACCGGAGGCGAGCAGTGGGGAATTTTGGACATGGGCGCA 351
DB      237799  TGAACACGGGCCAGACTCTTACCGGAGGCGAGCAGTGGGGAATTTTCCGCAATGGAGAA 237858

QY      352     AGCTGATCCAGCAATGCGCGTGAAGTGAAGAGGCTTGGGTTTGAAGCTCTTTGAG 411
DB      237859  AGTCTGAGGAGCAACCGCGGTGATGAAGAGGCTTTTCCGATCGTAAATACCTGTTGT 237918

QY      412     TCGAGAGAAAGGTTAGGTTAAATCGTGACCCATGAGGTTGACAGAGAAAGCA 471
DB      237919  TAGAGAAACAAAGGATAGAGTTAAGTCTGTTGCTTGAAGGTTATCAACAGAAACC 237978

QY      472     CCGGCTAATCTAGTCCAGACGCCGGTAAATACGTAGGTCGACGCTTAATCGGAAT 531
DB      237979  ACGGTTAATCTAGTCCAGACGCCGGTAAATAGTAGGCAAGCGTTGTCGGGATTT 238038

QY      532     ACTGGGGGTTAAAGGTGGCGAGGGCGCTTGAATCGATGATGAAATCCCGGGCTTAA 591
DB      238039  ATTGGGCGTTAAAGCGCGGCGAGGGCGCTTTTAAGTGAATGTGAAGAACCCCGGCTTAA 238098
```

```
QY      592     CTTGGGAATTTGGTTGAACCTTAAAGCTAGAGTGTGCGAGGAGGTGGATTCAT 651
DB      238099  CCGGGAGGGGTCAATGGAACTGGAACTGAGATGTCAGAAAGAGAGATGGATTCAC 238158

QY      652     GTGTAGCAGTGAATTCGTAGAGATATGGAAGAACATGATGCGAAGCGACTCTCGG 711
DB      238159  GTGTAGCGGTGAATTCGTAGATATGTGAGAGAACACAGTGGCGAAGCGACTCTCGG 238218

QY      712     GTTAACTGAGCGCTCATGACGAAAGCGTGGGAGCAACAGGATTTAATNCCCTGGTA 771
DB      238219  TCTGTAACTGAGCGCTGAGCGGCAAGCGTGGGAGCAACAGATTTAGTACCTGGTA 238278

QY      772     GTTCAAGCCCTTAAACGATGCACTAGTTGTT--GGGCGCTTATAGCTTGTGAAGAG 829
DB      238279  GTTCAAGCCGTAAACGATGCTAAGTGAAGGGGTTTCCGCCCTTATAGTGTGACAG 238338

QY      830     CTAAAGCGTGAAGTTGACCGCTGGGAGATACGGTGCAGAAATTAATAACTCAAGAAAT 889
DB      238339  CTAAAGCGTGAAGTTGACCTCGCTGGGAGATACCGGCAAGTTGAAGCTCAAGAAAT 238398

QY      890     GACGGGAGCCCGCAAGCGGTGATATGTGATTATTCGATGCAACCGGAAAAAAGCT 949
DB      238399  GACGGGAGCCCGCAAGCGGTGATATGTGATTATTCGATGCAACCGGAAAAAAGCT 238458

QY      950     TACTTACCTTGAATGATGACCAATTTTCTAGAGATGA-TTATGCTGTTGGGAAAGCTA 1008
DB      238459  TACCAGTCTTGAATGATGACCAATTTTCTAGAGATGA-TTATGCTGTTGGGAAAGCTA 238518

QY      1009  ACAAGGTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
DB      238519  TGAAGGTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 238578

QY      1069  ACGAGCGCAACCTTGTGATTAATGCGATCAATTTGTTGGGCACTTTAATGAGCTGCC 1128
DB      238579  ACGAGCGCAACCTTGTGATTAATGCGATCAATTTGTTGGGCACTTTAATGAGCTGCC 238638

QY      1129  GGTGACAAACCGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATG 1188
DB      238639  GGTG-CAAGCGGAGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATG 238697

QY      1189  CTTCAACGTTAATCAATGCGGCTGACAGAGGGTGGCAACCGCGAGGGGAGCTAATC 1248
DB      238698  CTTCAACGTTAATCAATGCGGCTGACAGAGGGTGGCAACCGCGAGGGGAGCTAATC 238757

QY      1249  TCAGAAAGCGCTGTGATGCTGCGATGCGAGTGGCACTGCACTCGCTGAAGTGGATC 1308
DB      238758  CCAATTAATCTATCTGATGCTGCGATGCGAGTGGCACTGCACTCGCTGAAGTGGATC 238817

QY      1309  GCTAGTATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1368
DB      238818  GCTAGTATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 238877

QY      1369  CCGTCAACCATATGAGATGGTTTCAACAAAGAGGTAGCTTAAC-GTAAAGAGGGCG 1427
DB      238878  CCGTCAACCATATGAGATGGTTTCAACAAAGAGGTAGGTAGCTTAATGAGAGCGAG 238937

QY      1428  CTTGCAAGGTGATGATTCATGACTGGGGTG 1457
DB      238938  CCGCCGAAGGTGGAGCAAGTATTTGGGGTG 238967

RESULT 64
US-11-045-004-1/c
Sequence 1, Application US/11045004
Publication No. US2006078901A1
GENERAL INFORMATION:
APPLICANT: BUCHRIESEN, CARMEN
APPLICANT: FRANGEUL, LIONEL
APPLICANT: COUVE, ELISABETH
APPLICANT: RUSNIOK, CHRISTOPHE
APPLICANT: FSIHI, HAFIDA
APPLICANT: DEHOIX, PIERRE
```

	APPLICANT:	DUSURGET, OLIVIER
	APPLICANT:	CHETOUANI, FARID
	APPLICANT:	NEDJARI, HAFED
	APPLICANT:	GLASER, PHILIPPE
	APPLICANT:	KUNST, FRANK
	APPLICANT:	COSSART, PASCALE
	APPLICANT:	DANIELS, JUSTIN
	APPLICANT:	GOSBELT, WERNER
	APPLICANT:	KREFT, JURGEN
	APPLICANT:	KUHNI, MICHAEL
	APPLICANT:	NG, EVA
	APPLICANT:	VÁZQUEZ-BOLAND, ANTONIO
	APPLICANT:	DOMÍNGUEZ-BERNAL, GUSTAVO
	APPLICANT:	GARRIDO-GARCÍA, PATRICIA
	APPLICANT:	TERRERRE-MARTINEZ, ALBERTO
	APPLICANT:	AMEND, ALEXANDRA
	APPLICANT:	CHAKRABORTY, TRINAD
	APPLICANT:	DOMANN, EUGEN
	APPLICANT:	HAIN, THORSTEN
	APPLICANT:	BECHE, PATRICK
	APPLICANT:	CHARBIT, ALAIN
	APPLICANT:	DURANT, LIONEL
	APPLICANT:	PÉREZ-DÍAZ, JOSÉ-CLAUDIO
	APPLICANT:	BAQUERO, FERNANDO
	APPLICANT:	GARCÍA DEL PORTILLO, FRANCISCO
	APPLICANT:	GÓMEZ-LOPEZ, NURIA
	APPLICANT:	MADUENO, ENCARNACIÓN
	APPLICANT:	PABLOS, BETRIZ DE
	APPLICANT:	WEHLAND, JURGEN
	APPLICANT:	KARST, UWE
	APPLICANT:	EHTIAN, KARL-DIETER
	APPLICANT:	HAUF, JORG
	APPLICANT:	ROSE, MATTHIAS
	APPLICANT:	VÖSS, HAMUT
	TITLE OF INVENTION:	LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
	FILE REFERENCE:	05394_0018-02
	CURRENT APPLICATION NUMBER:	US/11/045,004
	PRIOR FILING DATE:	2005-01-28
	PRIOR APPLICATION NUMBER:	10/637,657
	PRIOR FILING DATE:	2003-08-11
	PRIOR APPLICATION NUMBER:	10/257,023
	PRIOR FILING DATE:	2002-10-08
	PRIOR APPLICATION NUMBER:	PCT/FR01/01118
	PRIOR FILING DATE:	2001-04-11
	PRIOR APPLICATION NUMBER:	FR 00/04,629
	PRIOR FILING DATE:	2000-04-11
	NUMBER OF SEQ ID NOS:	2854
	SOFTWARE:	PatentIn version 3.3
	SEQ ID NO 1	
	LENGTH:	2944528
	TYPE:	DNA
	ORGANISM:	Listeria monocytogenes
	US-11-045-004-1	
	Query Match	56.6%; Score 824.8; DB 18; Length 2944528;
	Beet Local Similarity	76.3%; Pred. No. 4.3e-245;
	Matches 1121; Conservative	0; Mismatches 332; Indels 17; Gaps 8;
Oy	4	GAAGCGTCGGCCGACGCTTTACATGCAATGCGAAGCGCAGCACGCAGATGCTTGACATCT- 62
Db	1747787	GAACCGCTGCGCGCGCTTAATACATGCGAAGCGAAGCGAAGAAAGACTTGCTTT 1747728
Oy	63	-GGTGGCAGTAGCGCGAGCGGGTAGTAATGATCGG-AACGTATTCAGAAAGAGGGGGTA 120
Db	1747727	CCAAAGTTAGTGGCGGAGCGGGTAGTAACAGTGGCAACCTGCTGTAAATTGGGATA 1747668
Oy	121	ACGCATCGAAAGATGTGCTAATACGCAT-ATACTTAAGAGAGAAAGAGGGGATCGAA 179
Db	1747667	ACTCGGGAAACC GGCGCTTAATACGAATGATAGAATGTGGCGCATGCCGCTTTGAA 1747608
Oy	180	AGAC-----CTTGCGCTTTGGAGCGCGCCCATGTCTGATTAGCTAGTTGGTGGGTA 231
Db	1747607	AGATGGTTTCGGCTATCCCTTAACAAGATGGCGCGGATGCAATVGGCTAGTTGGTGGGTA 1747548

QY	232	AAGCCTTACCAAGCGCAGATCACTAGTGTCTGAGACGACCAACCTGGAC	291
Db	1747547	ATGCGCTTACCAAGCAACGATGCTATACCGACTGAAAGGATGATCGGCCACACTGGAC	1747488
QY	232	TGAGCAACGGCCGAGACTCTTACCGAGAGCGACGAGTGGGAAATTTTGGACAAATGGCGCA	351
Db	1747487	TGAGCAACGGCCGAGACTCTTACCGAGAGCGACGAGTGGGAAATTTTCCGCAATGGACGA	1747428
QY	352	AGCCTGATCCAGCAATGCGCGTGAAGTGAAGAGCGCTTCGGGTGTAAAGCTCTTCAAG	411
Db	1747427	AGTCTGACGAGCAACCGCGGTGTATGAAGAAAGTTTTCGATCTGTAAGTACTTGTGT	1747368
QY	412	TCGAGAGAAAAGGTTACGTTAAATTCGTGACCCATGACGCTATGACAGAAAGCA	471
Db	1747367	TAGAGAAAGAACAGATTAAGTAACTGCTTGTCCCTTTCGATTAACCAAGAGCC	1747308
QY	412	CCGGCTAACCTACGCGCAGCGCGCGTTAACTGTAAGGGGCAAGGCTTAAATGGAAT	531
Db	1747307	ACGGCTAACCTACGCGCAGCGCGCGTTAACTGTAAGGGGCAAGGCTTCCGAAATTT	1747248
QY	532	ACTGGGCGTAAAGGGTGGCAGCGCGCTTGTAAAGTGAATGTGAAATCCCGGGCTTAA	591
Db	1747247	ATTGGCGCTAAAGCGCGCGCGCTTGTAAAGTGAATGTGAAAGCCCGCGCTTAA	1747188
QY	592	CCTGGGAATTCGTTTGAACCTACAAAGCTAAGATGTGGCAGAGGAGGTGAATTCAT	651
Db	1747187	CCGGGAGAGGTCATTTGGAACCTGGAACCTGAGATGCAAGAGGAGATGGAATTCAC	1747128
QY	652	GTTGAGCACTGAATTCGTTAGATTAATGGAAGCAATCGATGGCGAAGGCAAGCTCTCG	711
Db	1747127	GTTGAGCGGTAAATTCGTTAATTTGAGAGAAACACAGTGGGAAAGCGAATCTCTCG	1747068
QY	712	GTTAAACATGACGCTCAGCAGAAACGTTGGGAGCAACAGATTAAGTATCCCTGTA	771
Db	1747067	TCTGTAATGACGCTGAGCGCGCAACCTGCGGAGCAACAGATTAAGTATCCCTGTA	1747008
QY	772	GTCACCGCCCTAAACGATGTCACCTAGTGT - GGGCTTAAATTAAGCTTGTGAACGA	829
Db	1747007	GTCACCGCGTAAACGATGATGCTAATGTTAGGGGGTTTCGCGCCCTTAATGTCCTAG	1746948
QY	830	CTAACCGCTGAATTTGACCCGCTGGGAGTACGCTCCCAAGATTAATACTCAAGAAAT	889
Db	1746947	CTAACCGCTTAAAGCACTCCGCTGGGAGTACGACCCCAAGTGAATCTCAAGAAAT	1746888
QY	890	GACGGGGACCGCGCAAGCGGTGATTAATGATTAATTCAGTGAACGCGAAGAACT	949
Db	1746887	GACGGGGACCGCGCAAGCGGTGAGCATGTGTTTAATTCAGAACGCGAAGAACT	1746828
QY	950	TACCTACCTTGACATGAGCAATTTTCTAGAGATAGA - TTAGTCTTCGGAGACGCTA	1008
Db	1746827	TACCAAGCTTGACATCTTCTTGAACACTCTGAGACAGAGCTTCCCTCGGGGACAAAG	1746768
QY	1009	AACAGGTGCTGATGCTGTCTGACGCTGTGCTGAGATGTTGGTTAATGCTCCGCA	1068
Db	1746767	TGACAGGTGCTGATGCTGTGCTGACGCTGTGCTGAGATGTTGGTTAATGCTCCGCA	1746708
QY	1069	ACGAGCGCAACCTTGATTAATTTGGCATATTTGGTGGGCACTTAAATGACACTGCC	1128
Db	1746707	ACGAGCGCAACCTTGATTAATTTGGCATTAATTTGGGCACTTAAATGACACTGCC	1746648
QY	1129	GATGCAAAACCGAGAGAGTGGGAGATGACCTCAAGTCTTCATGGACCTTAATGGGTAGG	1188
Db	1746647	GATG - CAAGCCGAGAGAGTGGGAGATGACCTCAAAATCATATGCTCCCTTAATGACCTGGG	1746589
QY	1189	CTTCAACGTAATCAATGGCGCTACAGAGGTTGCCAAACCGCGAGGGGAGCTAATC	1248
Db	1746588	CTAACAACGTCCTCAATGATAGTATGAACAAAGGTCGCGAAGCCGCGAGGTGAGACTAATC	1746529
QY	1249	TCAGAAAGCGGTGATGCTCCGATTCGAGATCGAGATCTGCAACTGCACTCCGTGAAGTCCGATC	1308
Db	1746528	CCATTAACCTATTTCTCAAGTTCGATTTGAGGCTTCAACTGCGCTTAACATGAAGCCGAGATC	1746469









Qy	1127	CCGGTGCACAAACCGGAGGAAGSTGGGATGACGTCACAGTCCCTCATTGAGCCCTTAATGGGTAG	1186
Db	1171	CCGGTGCACAAACCGAGAGGAAGSTGGGATGACGTCACAAATCAATGACCCCTTAATGATTTG	1230
Qy	1187	GGCTTTCACAGCTAATACATAGGCGCGTACACAGGGTTGCCAAACCCGCGAGGGAGNGTAA	1246
Db	1231	GGCTACACACGTGCTACATATGACACATACAAAGGCGAGCGAAACCGGAGGTCAAAGCAA	1290
Qy	1247	TCTCAGAAAGCGCGTCTGATGCCGATCGAGTCTGCATCTCGACTCCGTGAAGTCGAA	1306
Db	1291	TCCATTAAGTTGTTCTCAGTTCCGATTGTGATGTGCAACTCGACTCATGAACTGGAA	1350
Qy	1307	TGCGTAGTAATCGCGGATCAGCATGTCCGGGTGAATAGTTCGCCGGCTTGTGACACACC	1366
Db	1351	TCGCTAGTAATCGTAGATCACACATGCTCAACGATGAATAGTTCCGGGTATTTGTACACACC	1410
Qy	1367	GCCGTCACACATGCGAGTGGGTTTCACAGAAAGCAGTGTACTAAC-GTAAGAGAGG	1425
Db	1411	GCCGTCACACACAGAGTTTGTAAACCCGGAACCCGTTGGAGTAACCTTTTAGAGACT	1470
Qy	1426	CGCTTGACAGGTGAGATTCAATGACTGGGGTG	1457
Db	1471	AGCGTCGAAGTGGACAAATGATTTGGGGTG	1502

```

RESULT 67
US-11-198-794-160
; Sequence 160, Application US/11198794
; Publication No. US20060035257A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; APPLICANT: LYAMICHEV, VICTOR I.
; APPLICANT: OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; TITLE OF INVENTION: PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/198,794
; FILING DATE: 05-Aug-2005
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,193
; FILING DATE: 28-Aug-2001
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1555 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-11-198-794-160

```

Best Local Similarity 75.7%; Pred. No. 1.3e-245;  
Matches 115; Conservative 0; Mismatches 340; Indels 17; Gaps 7.

QY	3	TGAACGTGGCGGCATGCTTTACATATGCAAGTCGAACGGCAGACAG - GATGCTTGATC	61
Db	31	TGAACGTGGCGGCATGCTTTACATATGCAAGTCGAACGGCAGACAG - GATGCTTGATC	90
QY	62	T - GGTGGCAGATGGCGGACGGGTGAGTAATGCAATCGG - AACGTAATCAGAAAGAGGGGGT	119
Db	91	TCATGATTTAGCGGCGGACGGGTGAGTAACATGTAATACCTACCTATTAAGCTGGAT	150
QY	120	AACGCATCGAAAGATGTCTAATACCGCAT - -----TACTCTAAGAGAAAGCAGG	171
Db	151	AACCTTGGGAAACCGGAGCTAATACCGATAATATTTTGAACCGCAGTGTTCAAAGTGA	210
QY	172	GGATCGAAAGACCTTGGCGCTTTTGAAGCGGCCGATGTCTGATTAAGTAAGTTGGGGGTA	231
Db	211	AAGACGCTCTGTGCTGCTAATTATAGATGATTCGCGCTGCACTTAGCTAGTTGGTAAGTA	270
QY	232	AAGGCTTAACCAAGGCGACGATCAAGTAGTATGATCTGAGAGGACGACACCACTGGAC	291
Db	271	ACGGCTTAACCAAGGCGACGATCAAGTAGTATGATCTGAGAGGATGATGGCCACACTGGAC	330
QY	292	TGAGACACGGGCCGACCTCTTACGGAGGACGACAGTGGGAAATTTTGGCAATGGCGCA	351
Db	331	TGAGACACGGGTCCAGACTCTCTTACGGAGGACGACAGTGGGAAATTTTCCGCAATGGCGAA	390
QY	352	AGCCTGATCCAGCAATGCCCGGTGATGTAAGAAAGGCTTTGGGTTGTAAGCTCTTCA	411
Db	391	AGCCTGACGGGCAACCGCCGTGATGTAAGAAAGTCTTGGATCGTAATCTCTTAT	450
QY	412	TCGAGAAAGAAAGGTTACGGTAATATCGTGAACCCATGACGGTATGACAGAGAAACA	471
Db	451	TAGGAAAGAACTATGTGTATGATGTAACGTGCAACTTTTGAACGGTACTTAATCAGAAAGCC	510
QY	472	CCGGCTAATCACTGTCGACAGCGCGGTAAATACGTAGGCTGCAAGCTTAATGGAAT	531
Db	511	ACGGCTAATCACTGTCGACAGCGCGGTAAATACGTAGGCTGCAAGCTTAATCGGAAT	570
QY	532	ACTGGGCGTAAGAGGTGCGCAGAGGGCGCTTGTATGATGATGTGAATCCCGGCGTTAA	591
Db	571	ATTGGGCGTAAGAGGCGCGGTAGGCGGTTTTTAAAGTCTGATGTGAAGGCCACGCTCA	630
QY	592	CCTGGGAATTCGTTTGAACCTAACAAAGCTAGATGTCGACAGGAGGTGGAATTCAT	651
Db	631	CCGTGAGAGGTCATTTGGAACCTGGAATCTTGAATGACGAAAGGAAAGTGAATTCAT	690
QY	652	GTGTAGCAGTGAATGCGTAGAGATATGAAAGAAATCATGATGCGAAGCGACCTCTCG	711
Db	691	GTGTAGCAGTGAATGCGTAGAGATATGAAAGAAACATGATGCGAAGCGACCTCTCTCG	750
QY	712	GTTAACACTGACGCTCATGACGAAAGCGTGGGAGGAAACAGGATTAAGTATACCCTGTA	771
Db	751	TCTGTAACTGACGCTCATGATGTCGAAAGGCTGGGAGATCAACAGGATTAAGTATACCCTGTA	810
QY	772	GTCACGCGCCATAACGATGTCAACTAGTGT - -GGGCTTAATTAAGCTTGTGTAACGAG	829
Db	811	GTCACGCGCCATAACGATGTCAACTAGTGTTAAGGGGGTTTCGCGCCCTTAAGTGTCTGAG	870
QY	830	CTAACGCGCTGAAGTTTGACCGCTGGGAGTACGGTCCGACAGATTAAACTCAAGGAATT	889
Db	871	CTAACGCGCTGAAGCACTCGCGCTGGGAGTACGACCGCAAGTGTGAACCTCAAGGAATT	930
QY	890	GACGGGAGCCCGCACAAACGGGTGATTAATGTGATTTAATTCGATGCAACGGGAAAACT	949
Db	931	GACGGGAGCCCGCACAAACGGGTGAGCATGTGTGTTAATTCGAAGCAACGGGAAAGAACT	990
QY	950	TACCTAACCTTGAACATGTAGGAATTTTCTAGAGATAGA - -TTAGTGTCTTGGGAAAGC	1006
Db	991	TACCAATCTTGACATCTTTTGGAACAACCTTAGAGATAGAGCTTCCCTTGGGGGACAA	1056
QY	1007	TAAACAGATGCTGATGAGCTGTGCTCACTGCTGTGTGTAAGATTTGGCTTAAGTCCCG	1066

Db 1051 AGTGACAGGTGGTGCATGGTGTGTCCTGACGTGTCGTGCGTAGATGTTGGGTTAAGTCCCG 1110  
Qy 1067 CAACGACGCCAACCCCTTGTCTATTATTTGCCATCATTTGTTGGGCACCTTTAATGAGACTG 1126  
Db 1111 CAACGACGCCAACCCCTTAACTTAGCTTAGTTCATCATTAATGTTGGGCACCTCTAAGTTGACTG 1170  
Qy 1127 CCGGTGACAAACCCGAGGAAAGGTGGGATGAGTCGAAGTCCATGAGCCCTTATGGGTAG 1186  
Db 1171 CCGGTGACAAACCCGAGGAAAGGTGGGATGAGTCGAAGTCCATGAGTCCCTTATGATTTG 1230  
Qy 1187 GCGTTCACACGTATATCAATGCGCGCTGACAGAGGTTGCCAACCCCGAGGGGAGACTTA 1246  
Db 1231 GCGTTCACACGTGCTTCAATGACATTAAGAGGGAAGGAAACCGCGAGGTCAAGCATA 1290  
Qy 1247 TCTCAAAAACCGGTGCTAGTCCGATCGGATCGGAACTGCACTCCGCTGAAGTGGAA 1306  
Db 1291 TCCCAATAAGTTGTTCTCTCAGTTTCGATGTGATCTCAACTCGACTACATGAAGCTGGAA 1350  
Qy 1307 TGGCTATGATCGCGGATGACAGATGTCGGGTGAATACGTTCCGGGTCTGTGACACACC 1366  
Db 1351 TGGCTATGATCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1410  
Qy 1367 GCGGTTCACACCATGAGGAGTGGGTTTACCAGAGAGGAGTGTAACTGTAAGAGAGG 1425  
Db 1411 GCGGTTCACACCATGAGGAGTGGGTTTACCAGAGAGGAGTGTAACTGTAAGAGAGG 1470  
Qy 1426 CCGTTCACACCATGAGGAGTGGGTTTACCAGAGAGGAGTGTAACTGTAAGAGAGG 1457  
Db 1471 AGCGGTCAAGAGGTGGACAAATGATGGGGTG 1502

RESULT 68  
US-11-198-657-160  
; Sequence 160, Application US/11198657  
; Publication No. US20060040299A1  
; GENERAL INFORMATION:  
; APPLICANT: BROW, MARY ANN D.  
; APPLICANT: LYAMICHEV, VICTOR I.  
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
; TITLE OF INVENTION: PATHOGENS  
; NUMBER OF SEQUENCES: 165  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/11/198,657  
; FILING DATE: 05-Aug-2005  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US/09/941,193  
; FILING DATE: 28-Aug-2001  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: FORS-01756  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 160:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1555 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double

; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-11-198-657-160  
Query Match 56.6%; Score 824; DB 18; Length 1555;  
Best Local Similarity 75.7%; Pred. No. 1.3e-245;  
Matches 115; Conservative 0; Mismatches 340; Indels 17; Gaps 7;  
Qy 3 TGAAGGCTGGCGGACGCTTTCATACATGCAAGTGCAGGCGACAG-GATGCTTCATC 61  
Db 31 TGAAGGCTGGCGGACGCTTTCATACATGCAAGTGCAGGCGACAGGAGGCTTC 90  
Qy 62 T-GTGGCGAGTGGCGGACGCGGTGATATGATCGG-AACTATCCAGAGAGGGGT 119  
Db 91 TGTGATGTTAGGGGAGGAGCGGTGATGATCAAGTGGATTAACCTATTAAGACTGGGAT 150  
Qy 120 AAGCATGGAAGATGTGCTAATACCGCAT------TACTTAAGAGAGAAAGCAG 171  
Db 151 AACTTCGGGAAACCGGAGCTAATACCGGATATATTTGAAACCGCATGGTTCAAAAGTGA 210  
Qy 172 GGATGGAAGAACCTTGCCTTTTGGAGGCGCGATGTGATTAAGTATGATGTTGGTGGTGA 231  
Db 211 AAGAGGCTTGTGCTGTCACTTATGATGATGATGATGATGATGATGATGATGATGATGAT 270  
Qy 232 AAGGCTTACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 291  
Db 271 AAGGCTTACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 330  
Qy 292 TGAAGACGCGCCGACGCTTCATGCGGAGGCGACGCTGGGAAATTTTGGCAATGGCGCA 351  
Db 331 TGAAGACGCGCCGACGCTTCATGCGGAGGCGACGCTGGGAAATTTTGGCAATGGCGCA 390  
Qy 352 AAGGCTTACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 411  
Db 391 AAGGCTTACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450  
Qy 412 TCGAGAGAAAGGTTACGCTTAATATCTGATGATGATGATGATGATGATGATGATGATGATGAT 471  
Db 451 TCGAGAGAAAGGTTACGCTTAATATCTGATGATGATGATGATGATGATGATGATGATGATGAT 510  
Qy 472 CCGGCTTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531  
Db 511 ACGGCTTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570  
Qy 532 ACTGGGCGTAAAGGTTGGCGGACGCGCTTGAATGATGATGATGATGATGATGATGATGATGAT 591  
Db 571 ACTGGGCGTAAAGGTTGGCGGACGCGCTTGAATGATGATGATGATGATGATGATGATGATGAT 630  
Qy 592 CCGGGAATTTGGCTTGAATACCAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 651  
Db 631 CCGGGAATTTGGCTTGAATACCAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 690  
Qy 652 GTGTGACATGAAATGCGTAGATATGGAAGAACATGATGCGGAGCGGACGCTCTG 711  
Db 691 GTGTGACATGAAATGCGTAGATATGGAAGAACATGATGCGGAGCGGACGCTCTG 750  
Qy 712 GTTAACTGACGCTTCAATCAAGAAAGCTGGGAGCAACAGATTAGATACCCCGGTA 771  
Db 751 TCTGTAACGACGCTTCAATCAAGAAAGCTGGGAGCAACAGATTAGATACCCCGGTA 810  
Qy 772 GTTCAAGCCCTTAAAGATGATCAATGATGTTGTT--GGGCTTATTAAGGCTTGTGTAAGCAAG 829  
Db 811 GTTCAAGCCCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870  
Qy 830 CTAAAGCGTGAAGTTGACCGGCTGGGAGTACGCTGCAAGATTAATACTCAAGAGAT 889  
Db 871 CTAAAGCGTGAAGTTGACCGGCTGGGAGTACGCTGCAAGATTAATACTCAAGAGAT 930  
Qy 890 GACGGGACCCGCAAGAGGCTGATTAATGATGATTAATTCATGCAAGCGGAAAGAACT 949  
Db 931 GACGGGACCCGCAAGAGGCTGATTAATGATGATTAATTCATGCAAGCGGAAAGAACT 990  
Qy 950 TACCTACCTTGAACATGATGAGCAATTTCTAGAGATAGA--TTAGTGTCTGGGAAACCG 1006

```
Db 991 TACCAATCTTGAACCTCTTGGACAACTTAGAGATAGACCTTCCCTTGGGGGACAA 1050
Qy 1007 TAAACAGAGTGCATAGCTGCTGTCAGCTCGTGTGAGATGTTGGTAAAGTCCG 1066
Db 1051 AGTGAACAGGTGTCATGATGTTGTCAGCTCGTGTGAGATGTTGGTAAAGTCCG 1110
Qy 1067 CAACGAGCGCAACCTTGTCTAATTAATTCATCATTTGTTGGCACTTTAATGAGACTG 1126
Db 1111 CAACGAGCGCAACCTTAACTTAGTTCATTAATTAATTTGGGCACTTAATGAGACTG 1170
Qy 1127 CCGGTGACAAACCGGAGAAAGTGGGAGTACGTCAGAGTCTCATGGCCCTTAATGGTAG 1186
Db 1171 CCGGTGACAAACCGGAGAAAGTGGGAGTACGTCAGAGTCTCATGGCCCTTAATGATTG 1230
Qy 1187 GGGTCAACGTAATACATAGGCGGTACAGAGGGTGGCCAAACCGGAGGGGAGCTAA 1246
Db 1231 GGGTCAACGTCATGTCATACATAGGCAATACAAAGGGCAGCAACCGGAGGTCAAGCAA 1290
Qy 1247 TCTCAGAAACGCGCTCGTAGTCCGAGTCGAGTCTGCACTCGATCGTGAAGTCCGAA 1306
Db 1291 TCCCATTAAGTGTGTCAGTTCGAGTTGATGTCGCACTCGACTACATGAAGCTGAA 1350
Qy 1307 TCGGTAGTAATCGCGGATCAGCATGTGCGGTGAATACGTTCCGGGTCTTTGACACAC 1366
Db 1351 TCGGTAGTAATCGTAGATCAGCATGTGCGGTGAATACGTTCCGGGTATTTGACACAC 1410
Qy 1367 GCGGTGACACCAATGGAGTGGGTTTCCAGGAGAGAGTACTTAAC-CTAAGAGAGG 1425
Db 1411 GCGGTGACACCAAGAGTGTGTAACACCGAGCCGGTGGAGTAACTTTTAGAGACT 1470
Qy 1426 CGCTTGCACGCTGAGATTCATGACTGGGGTG 1457
Db 1471 AGCGTGAAGTGGGCAAAATGATGGGGTG 1502
```

## RESULT 69

```
US-11-151-847-1
; Sequence 1, Application US/11151847
; Publication No. US2006002909A1
; GENERAL INFORMATION:
; APPLICANT: Kabushiki Kaisha Yonezawa Biru Shisutemu Sabisu
; TITLE OF INVENTION: A microorganism separated from Kefir grains, a microorganism cult
; TITLE OF INVENTION: obtained by culturing this microorganism or microorganisms inclu
; TITLE OF INVENTION: and a product using them.
; FILE REFERENCE: H16P1032
; CURRENT APPLICATION NUMBER: US/11/151,847
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 1521
; TYPE: DNA
; ORGANISM: Lactobacillus mali
US-11-151-847-1
```

```
Query Match 56.2%; Score 818.8; DB 17; Length 1521;
Best Local Similarity 76.3%; Pred. No. 5.2e-244;
Matches 1070; Conservative 1; Mismatches 318; Indels 14; Gaps 5;
```

```
Qy 69 GAGTGGCGGACGGGTGAGTAATGCTCGG-AACTGATCAAGAGGGGGGTAAAGCATC 127
Db 100 GAGTGGCGAAGCGGTGAGTAACACGTGGTAACTGCCCAAAAGAGGGGATTAACATTTG 159
Qy 128 GAAAGATGTGTAATACCGCATATCTCTAAGAGAGAAAGC-----AGGGATCGA 178
Db 160 GAAACAGGTGCTAATACCGCATACAAACAAACCGCTGTTTGTTTAAAGATGCT 219
Qy 179 AAGACTTGGCGCTTTTGGAGCGCGATGTCTGATTAGCTAGTTGGTGGGTAAGGCT 238
Db 220 TTGGCTATCACTTTTGGATGAGACCGGGCGGTATTAAGTGTGTAAGGATGGCTT 279
Qy 239 ACCAAGGCGAGCATGATGTTGTTCTAGAGAGACCAAGCCACACTGGGATGAGACA 298
```

```
Db 280 ACCAAGGCGATGATACGTAGCCGAACGTAGAGAGTTGATCGGCCACATGGGACTGAGACA 339
Qy 299 CGGCCCAAGTCTTACGGAGGCGAGAGTGGGAAATTTTGGACAAATGGGGCAAGCTGA 358
Db 340 CGGCCCAAGTCTTACGGAGGCGAGAGTGGGAAATTTTGGACAAATGGGCAAGCTGA 399
Qy 359 TCCAGCAATGCGCGTGAAGTGAAGAGGCTTGGGGTTGAAAGCTTTCACTGACGAGAA 418
Db 400 TGAAGCAAGCGCGGTGATGAGAAAGTTTTCGATCGTAAACTCTGTTGTAAGAA 459
Qy 419 GAAAGTTACGTAAATATCTGACCCATGACGGTATGACAGAAAGCAACCGCTTA 478
Db 460 GAAAGTTACGTAAATATCTGATCATRCAAGTACGGTATCTAAGAAAGCAACCGCTTA 519
Qy 479 ACTAAGTCCAGACCGCGGTAAATACGTAGGGTCAAGGTTAAATCCGAAATTAAGGGC 538
Db 520 ACTAAGTCCAGACCGCGGTAAATACGTAGGGTCAAGGTTAAATCCGAAATTAAGGGC 579
Qy 539 GTAAAGGGTGGCGAGCGGCTTGTAAATGATGATGAAATCCCGGGCTTAACTGGGA 598
Db 580 GTAAAGGGTGGCGAGCGGCTTGTAAATGATGATGAAATCCCGGGCTTAACTGGGA 639
Qy 599 ATTGCGTTGAAACTCAAGAGTGAAGTGGGAGAGGAGTGAATTCATGTATAC 658
Db 640 CGTGAATGGAACCTGGAGACCTTGAAGTGAAGAGAGAGTGAATTCATGTATAC 699
Qy 659 AGTGAATGCGTGAAGATATGAAAGCAATCATGATGGCGAAGGCACTCTGGGTTAA 718
Db 700 GGTGAATGCGTGAAGATATGAAAGCAATCATGATGGCGAAGGCACTCTGGGTTAA 759
Qy 719 CTGACCGTCAATGACAGAAACGTGGGGAGCAACAGGATTAGATACCTGTAGTCCAGC 778
Db 760 CTGACCGTCAATGACAGAAACGTGGGGAGCAACAGGATTAGATACCTGTAGTCCAGC 819
Qy 779 CCTTAAACGATGTCATAGTGTGGGCTTATTAAGG--TTGTAAAGAACTAAAGC 836
Db 820 CTGTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 879
Qy 837 GTGAAGTTGACCGCTGGGAGATACGCTCGCAGATTAATACTTAAGAAATTAAGCGGG 896
Db 880 ATTAAGCATTCGCGCTGGGAGATACGACCCCAAGGTTGAATCTCAAGAAATTAAGCGGG 939
Qy 897 ACCCGCAACGCGGTGATATGAGTAATTCGATGACAGCGGAAACCTTAACCTAC 956
Db 940 GCGCGCAACGCGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 999
Qy 957 CTTGACATGATGAGGAAATTTCTAGAGAT-AGATTAGTGTCTTGGGAAACGTAAACAGG 1015
Db 1000 TCTTGACATCTTTTGTCTAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1059
Qy 1016 TGCTGATGCTGTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
Db 1060 TGCTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1119
Qy 1076 CAACCTTGTCTAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1135
Db 1120 CAACCTTGTCTAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1179
Qy 1136 AACCGAGAAAGTGGGAGTGAACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACA 1195
Db 1180 AACCGAGAAAGTGGGAGTGAACGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTACA 1239
Qy 1196 CGTAATACATGAGCGGTACAGAGGTTGCCAACCCCGGAGGGGAGTAAATCTCAAGAA 1255
Db 1240 CGTGTCAATGACGATACAGAGTCCGAAAGCCGAGGATTAAGTAACTCTTAA 1299
Qy 1256 GCGGTGTATCCCGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1315
Db 1300 ACCGTTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
Qy 1316 ATCGGATCAGCATGTCGCGGTAAATAGTTCGCGGTCTTGTAAACACCGCGCTGAC 1375
Db 1360 ATCGGATCAGCATGTCGCGGTAAATAGTTCGCGGTCTTGTAAACACCGCGCTGAC 1419
```

```

RESULT 70
US-11-273-617-3
; Sequence 3, Application US/11273617
; Publication No. US20060067924A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Margie D.
; APPLICANT: Hofacres, Charles
; APPLICANT: Harmon, Barry
; TITLE OF INVENTION: Probiotic Bacteria and Methods
; FILE REFERENCE: 18-03
; CURRENT APPLICATION NUMBER: US/11/273,617
; CURRENT FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/470,807
; PRIOR FILING DATE: 2003-05-14
; PRIOR APPLICATION NUMBER: WO us2004/015378
; PRIOR FILING DATE: 2005-04-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Lactobacillus crispatus
US-11-273-617-3

```

Query Match	56.0%;	Score 816.6;	DB 18;	Length 1510;
Best Local Similarity	75.7%;	Pred. No. 2.5e-243;		
Matches 1066;	Conservative	0;	Mismatches 329;	Indels 13; Gaps 4

QY	63	GGTGGCAGATGGCGGACGGGGTGAGTAATGCATCTGG-AAAGTAATCCAGAAAGGGGGGTAA	121
Db	80	GGAAAGCAGACGGCGGATGGGTAGTAACACGTGGGTAACTGCTTAAGTCTGGGATAC	139
QY	122	CGCATCGAAAGATGTGCTAATAACGCATATTACTTAAGGAGGAAAGCAGGGATCGAAA-	180
Db	140	CATTGGAAAACAGGTGCTAATACCGGATTAACAACATGTATCGCATGATCGATGTTGAAA	199
QY	181	-----GACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTAA	232
Db	200	GGCGGCGTAAGCTGTGCTGCTAAAGGATGACCCGCGGTGATTAGCTAGTTGGTAAAGTTA	259
QY	233	AGGCGTAACCAAGGCGACGATCAGTAGTGTGCTGAGAGGACGACGACCACTGGGACT	292
Db	260	CGGCTTACCAAGGCAACGATGCATACCGAGTTGAGAGCTGATCGGCCACATTGGGACT	319
QY	293	GAGACACGGCGCCGACTCTCTACGGGAGGCGACGAGTGGGAAATTTTGCACATGCGCGCAA	352
Db	320	GAGACACGGCCCAACTCTCTACGGGAGGCGACGATGAGCAATCTTCACAATGGCGAAA	379
QY	353	GCTTATCCAGCAATGCGCGCTGATGTAAGAACGCTTGGGTTGTAAGCTCTTCACT	412
Db	380	GCTTATGAGCAACGCGCGCTGAGTGAAGAGGTTTTCGATCGTAAGGCTCTGTGTT	439
QY	413	CGAGAAAGAAAGTTACGTTAATAATCTGACCCATGACGGATTCGACAGAAAGAAC	472
Db	440	GGTGAAGAAAGATAGAGGTAGTAATCTGGCTTTTATTTGACGGTAATCAACAGAAAGTCA	499
QY	473	CGGCTAACTACCTGCGACAGCGCGGCTAATACGTAGGTGCAAGCGTTAATCGAATTA	532
Db	500	CGGCTAACTACCTGCCAGACGCGCGGTAAATGTAAGTGGGCAAGGTTGTGCGGATTTA	559
QY	533	CTGGGCGTAAAGGTTGCGCAGCGCGCTTGTAAATCAGATGTGAAAATCCCGGGCTTAAC	592
Db	560	TTGGGCGTAAACCGAGCGCGGAGGAAATTAAGTCTGATGTGAAAGCCCTCGGCTTAAC	619

QY	533	CTGGGAATTCGCTTGGAAACTACAAAGCTAGAGTGTGCGAGAGGGAGGTGGAATTCATG	652
Db	620	CGGGAAATGTACATCGGAACTGTTTTCTTGAGTCCAGAGAGGAGAGTGGAACTTCATG	679
QY	653	TGTAGCACTGAAATATGCTGAGATATGGAAGAACATCGATGGCGAAGGCAAGCCTCTGSG	712
Db	680	TGTAGCGGTGGAAATGCTGATGATATATGGAAAGAACACAGTGGCGAAGGCGGCTCTGTGT	739
QY	713	TTAACACTGAGCGCTATGACAGAAAGCGTGGGGAGCCAAACAGGATTTAGTATACCTTGTTG	772
Db	740	CTGCAACTGAGCGCTGAGGCTCGAAAGCATGGGTATCGAACAGAGATTTAGTATACCTTGTTG	799
QY	773	TCCAAGCCCTTAAACGATGTCACATGATGTTGGGCTTATATGAGCTTG--GTAAAGAAAGC	830
Db	800	TCATATCCGTTAAACGATGATGCTTAAGTGTGGAGAGTTCCGCTCTCAAGTCTGCAAC	859
QY	831	TAAACGCTGAAGTTGACCGGCTGGGAGATACGGTGCACAGATTTAAACTCAAGAAATTTG	890
Db	860	TAAACGATTTAAGCACTCCGCTGGGAGATACACCGCAAGGTTGAAACTCAAGGAATTTG	919
QY	891	ACGGGAGACCCCAACAAAGGGTGGATTTAATGTGATTTCAATGCAATGCAACGCGAAAAACCTT	950
Db	920	ACGGGGGCCCCGACAAAGCGGTGACATGTGTGTTAATTCGAAGCAACCGAAGAACTTT	979
QY	951	ACCTTACCCCTTGACATGTAGCGCAATTTTCTAGAGATA-GATTAAGTCTTGGGAGCGCTAA	1000
Db	980	ACCAAGTCTTGACATCTAGCGCAATTTGTAAGATATACGAAGTTTCCCTTGGGGAGCGCTAA	1030
QY	1010	CACAGATGCTGATGAGCTGTGCTGACAGTCTGCTCCTGAGATGTTGGTTAACTCCCGCA	1060
Db	1040	GACAGGTGTGATGAGCTGTGCTGACAGTCTGCTGAGATGTTGGTTAACTCCCGCA	1090
QY	1070	CGAGCGCAACCTTGTCAATTAATTCACATCATTTGTTGGGCACTTTAATGACCTGCCG	1120
Db	1100	CGAGCGCAACCTTGTCAATTAATTCACATCATTTGTTGGGCACTTTAATGACCTGCCG	1150
QY	1130	GTGACAAACCGGAGGAGGTGGGGATGACGTCAAGTCTCATAGGCCCTTAATGGGTAGGGC	1180
Db	1160	GTGACAAACCGGAGGAGGTGGGGAGACGTCAAGTCTCATAGGCCCTTATATACCTTGAGC	1210
QY	1190	TTCAACAGCTAATACAAATGAGCGGCTGACAGAGGGTTTCCAAACCGCGAGGGGAGCTAATCT	1240
Db	1220	TACACAGCTGTCAACATGGGCACTACACAGAGAACAAACCTGCGAAGCGAACGAAATCT	1270
QY	1250	CAGAAAGCGCGTGTGATGTCGGGATCGGAGTCTGCACTTGACTCCGTGAAGTGGGAATG	1300
Db	1280	CTGAAGCTGTTCCTCAATTCGGACTGTAGGCTGTCAACTCGCCTTAACGAAAGCTGGAAATG	1330
QY	1310	CTAGTAATCGGGGATCACAGATGTGCGGGGTAAATAGTTCCGGGGCTTGTATACACCGGC	1360
Db	1340	CTAGTAATCGGGGATCACAGATGTGCGGGGTAAATAGTTCCGGGGCTTGTATACACCGGC	1390
QY	1370	CGTCAACCATGAGAGTGGGTTTTCACGAAGCAGATGTTTAAACCGTAAAGAGGCGGCT	1420
Db	1400	CGTCAACCATGAGAGTCTGCAAGTCCCAAGCGCGGTGCTTAACCTTGGGAAGGAGGC	1450
QY	1430	TGCCACGGTGAATTCATGACTGGGGTG 1457	
Db	1460	GTCTTAAGCAGGGCAGATGACTGGGGTG 1487	

Search completed: May 5, 2006, 14:35:41  
Job time : 791.5 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceeleration Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 5, 2006, 13:25:44 ; Search time 755.5 Seconds  
(without alignments)  
7853.986 Million cell updates/sec

Title: US-10-659-980A-1  
Perfect score: 1457  
Sequence: 1 atgaacgctgcgcgcacatgc.....tgagatcatgactcggggtc 1457

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 70 summaries

Database : Published Applications NA New:\*  
1: /SID55/ptodaca/1/pubpna/US08\_NEW\_PUB.seq1:\*  
2: /SID55/ptodaca/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /SID55/ptodaca/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /SID55/ptodaca/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /SID55/ptodaca/1/pubpna/PCT\_NEW\_PUB.seq:\*  
6: /SID55/ptodaca/1/pubpna/US09\_NEW\_PUB.seq:\*  
7: /SID55/ptodaca/1/pubpna/US09\_NEW\_PUB.seq2:\*  
8: /SID55/ptodaca/1/pubpna/US09\_NEW\_PUB.seq2:\*  
9: /SID55/ptodaca/1/pubpna/US10\_NEW\_PUB.seq1:\*  
10: /SID55/ptodaca/1/pubpna/US10\_NEW\_PUB.seq2:\*  
11: /SID55/ptodaca/1/pubpna/US10\_NEW\_PUB.seq2:\*  
12: /SID55/ptodaca/1/pubpna/US10\_NEW\_PUB.seq4:\*  
13: /SID55/ptodaca/1/pubpna/US11\_NEW\_PUB.seq:\*  
14: /SID55/ptodaca/1/pubpna/US11\_NEW\_PUB.seq2:\*  
15: /SID55/ptodaca/1/pubpna/US11\_NEW\_PUB.seq2:\*  
16: /SID55/ptodaca/1/pubpna/US11\_NEW\_PUB.seq4:\*  
17: /SID55/ptodaca/1/pubpna/US11\_NEW\_PUB.seq5:\*  
18: /SID55/ptodaca/1/pubpna/US11\_NEW\_PUB.seq5:\*  
19: /SID55/ptodaca/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1225.4	84.1	1508	US-11-248-383-1	Sequence 1, Appl1
2	1149.2	78.9	1522	US-11-198-242-2	Sequence 2, Appl1
3	1142.8	78.4	1526	US-11-198-242-1	Sequence 1, Appl1
4	1089.6	74.8	1405	US-10-515-311-5	Sequence 5, Appl1
5	1021.8	70.1	1501	US-10-513-639-1	Sequence 1, Appl1
6	1021.2	70.1	1493	US-10-831-286A-48683	Sequence 48683, A
7	1020.2	70.0	1510	US-10-831-286A-48685	Sequence 48685, A
8	1019.8	70.0	1510	US-10-831-286A-48680	Sequence 48680, A
9	1019.6	69.9	1507	US-10-831-286A-48681	Sequence 48681, A
10	1018	69.9	1507	US-10-831-286A-48687	Sequence 48687, A
11	1017	69.8	1508	US-10-831-286A-48684	Sequence 48684, A
12	1015.4	69.7	1510	US-10-831-286A-48682	Sequence 48682, A
13	1008.4	69.2	1530	US-10-831-286A-48671	Sequence 48671, A
14	1007.2	69.1	1492	US-10-831-286A-48686	Sequence 48686, A

## ALIGNMENTS

15	1007.2	69.1	1519	US-10-831-286A-48687	Sequence 48687, A
16	1005.4	69.0	1495	US-10-831-286A-48675	Sequence 48675, A
17	1005.2	69.0	1494	US-10-831-286A-48669	Sequence 48669, A
18	1005.2	69.0	1503	US-10-831-286A-48665	Sequence 48665, A
19	1005.2	69.0	1514	US-10-831-286A-48666	Sequence 48666, A
20	1004.8	69.0	1494	US-10-831-286A-48676	Sequence 48676, A
21	1003.8	68.9	1535	US-10-831-286A-48670	Sequence 48670, A
22	1003.6	68.9	1501	US-10-831-286A-48678	Sequence 48678, A
23	1002.8	68.8	1527	US-10-831-286A-48693	Sequence 48693, A
24	1000	68.6	1486	US-10-831-286A-48679	Sequence 48679, A
25	995.2	68.3	1542	US-09-941-095-158	Sequence 158, App
26	995.2	68.3	1542	US-11-198-746-158	Sequence 158, App
27	995.2	68.3	1542	US-11-198-794-158	Sequence 158, App
28	995.2	68.3	1542	US-11-198-657-158	Sequence 158, App
29	990	67.9	1533	US-10-831-286A-48688	Sequence 48688, A
30	988	67.8	1418	US-10-831-286A-48691	Sequence 48691, A
31	988	67.8	1418	US-10-831-286A-48692	Sequence 48692, A
32	981.2	67.3	1436	US-10-831-286A-48667	Sequence 48667, A
33	979.8	67.2	1481	US-10-831-286A-48672	Sequence 48672, A
34	977.8	67.1	1525	US-10-831-286A-48689	Sequence 48689, A
35	850.4	58.4	1485	US-11-055-637-76	Sequence 76, Appl
36	848.8	58.3	1513	US-11-055-637-79	Sequence 79, Appl
37	845.8	58.1	1505	US-11-055-637-80	Sequence 80, Appl
38	841.2	57.7	1509	US-11-273-617-6	Sequence 69, Appl
39	839.6	57.6	1486	US-11-055-637-69	Sequence 66, Appl
40	838.6	57.6	1545	US-11-055-637-66	Sequence 8, Appl
41	838.6	57.6	1587	US-11-273-617-8	Sequence 71, Appl
42	837.2	57.5	1486	US-11-055-637-71	Sequence 78, Appl
43	837.2	57.5	1507	US-11-055-637-78	Sequence 3356, Ap
44	837	57.4	3169	US-10-793-626-3356	Sequence 3905, Ap
45	832	57.1	1433	US-10-793-626-187	Sequence 4187, Ap
46	837	57.4	3657	US-10-793-626-4460	Sequence 3955, Ap
47	833.8	57.2	3008	US-10-793-626-4460	Sequence 3955, Ap
48	833.8	57.2	3821	US-10-793-626-3955	Sequence 3957, Ap
49	833.8	57.2	3821	US-10-793-626-3955	Sequence 3957, Ap
50	832.4	57.1	1510	US-11-273-617-9	Sequence 9, Appl1
51	832	57.1	1408	US-10-953-392-1	Sequence 1, Appl1
52	832	57.1	1408	US-10-954-147-1	Sequence 15, Appl1
53	832	57.1	1433	US-10-953-392-15	Sequence 15, Appl1
54	832	57.1	1433	US-10-954-147-15	Sequence 15, Appl1
55	831.8	57.1	1470	US-11-055-637-84	Sequence 84, Appl1
56	830.8	57.0	1368	US-10-953-392-10	Sequence 10, Appl1
57	830.8	57.0	1368	US-10-954-147-10	Sequence 83, Appl1
58	828.6	56.9	1491	US-11-055-637-83	Sequence 4, Appl1
59	827.2	56.8	1395	US-10-953-392-4	Sequence 74, Appl1
60	827.2	56.8	1395	US-10-953-392-4	Sequence 74, Appl1
61	827	56.7	1555	US-09-941-095-160	Sequence 160, App
62	825.6	56.7	1555	US-11-198-746-160	Sequence 160, App
63	825.6	56.7	1555	US-11-198-746-160	Sequence 160, App
64	825.6	56.7	1555	US-11-198-657-160	Sequence 160, App
65	825.6	56.7	1555	US-11-198-657-160	Sequence 160, App
66	823.4	56.5	1459	US-11-084-508-3	Sequence 3, Appl1
67	823.2	56.5	2944528	US-11-045-004-1	Sequence 1, Appl1
68	823.2	56.5	2944528	US-11-045-004-1	Sequence 1, Appl1
69	817.2	56.1	1521	US-11-151-847-1	Sequence 1, Appl1
70	816.6	56.0	1510	US-11-273-617-3	Sequence 3, Appl1

RESULT 1  
US-11-248-383-1  
; Sequence 1, Application US/11248383  
; Publication No. US20060081532A1  
; GENERAL INFORMATION:  
; APPLICANT: Draheo, David J.  
; TITLE OF INVENTION: Consortium of Nitirifying Bacteria  
; FILE REFERENCE: 10708-200-US  
; CURRENT APPLICATION NUMBER: US/11/248,383  
; CURRENT FILING DATE: 2005-10-11  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.3

SEQ ID NO 1  
; LENGTH: 1508  
; TYPE: DNA  
; ORGANISM: Nitrosomonas eutropha  
US-11-248-383-1

Query: March 84.1%; Score 1225.4; DB 18; Length 1508;  
Best Local Similarity 92.1%; Pred. No. 0; Mismatches 106; Indels 9; Gaps 5;  
Matches 1348; Conservative 0;

```
Oy 1 ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGCAGGATCTTGAT 60
Db 24 ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAACGGCAG--CGGGGCTTCGGC 81
Oy 61 CTGTGGCGAGTGGCGGACGGGTGAGTAATGATCGGAACGTATCCAGAAAGGGGGGTA 120
Db 82 CTGGCGGCGAGTGGCGGAACGGGTGAGTAATACATCGGAACGTGCTTGAGTGGGGAATA 141
Oy 121 ACGCATCGAAAGATGCTTAATACCGCATATATCTTAAGGAGGAAGAGCGGATCGAAA 180
Db 142 ACGCATCGAAAGATGCTTAATACCGCATAT-TTCTCGAAAGAAAGAGCGGATCGAAA 200
Oy 181 GACCTTGCGCTTTTGGAGCGGCGCATGCTGATTAAGTGTGAGGAGTAAAGGCTTAC 240
Db 201 GACCTTGCGCTTAAGAGAGCGGCGCATGCTGATTAAGTGTGAGGATTAAGGCTTAC 260
Oy 241 CAAGCGCATGATCAATGTTGGTCTGAGAGGACGACGACCACTGGGACTGAGACACG 300
Db 261 CAAGCGCATGATCAATGCTGCTGAGAGGACGATCAGCCACACTGGGACTGAGACACG 320
Oy 301 GCCCGACTCCTTACGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 321 GCCCGACTCCTTACGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 380
Oy 361 CAGCAATCCCGGCTGAGTGAAGAAAGGCGCTTCGGGTGTAAGGCTTTTCACTGAGAA 420
Db 381 CAGCAATCCCGGCTGAGTGAAGAAAGGCGCTTCGGGTGTAAGGCTTTTCACTGAGAA 440
Oy 421 AAAGTTACGGTAATTAATCTGACTCATGACGGTATGACAGAAAGAACACCGGCTTAC 480
Db 441 AAAGTTATGTTAATTAATCAATGATTTATGACGGTATGACAGAAAGAACACCGGCTTAC 500
Oy 481 TACCGTCAGACGCGCGGCTTAATGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Db 501 TACCGTCAGACGCGCGGCTTAATGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 560
Oy 541 AAAGGATCGGACGCGGCTTTGTAGTCAATGATGTAATCCCGGGCTTAACTTGGGAT 600
Db 561 AAAGGATCGGACGCGGCTTTGTAGTCAATGATGTAATCCCGGGCTTAACTTGGGAT 620
Oy 601 TCGCTTTGAAACTACAGAGCTAAGTGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 621 TCGCTTTGAAACTACAGAGCTAAGTGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGG 680
Oy 661 TGAATTCGGTGAAGATTAAGAAAGCAATCGATGGGAGGAGGAGGAGGAGGAGGAGG 720
Db 681 TGAATTCGGTGAAGATTAAGAAAGCAATCGATGGGAGGAGGAGGAGGAGGAGGAGG 740
Oy 721 GACGCTCATGACGAAACCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 741 GACGCTCATGACGAAACCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 800
Oy 781 CTAAACGATGCTCACTAGTTGTTGGGCTTATTA--GGCTTGTATACGAAAGCTAACGCGT 838
Db 801 CTAAACGATGCTCACTAGTTGTTGGGCTTATTAAGATTGTTGTTAAGCTTAACTAACGCGT 860
Oy 839 GAAGTTGACCGGCTGGGAGGAGTACCGTCCGAAAGATTAAATCCCAAGGAATTGACGGGAC 898
Db 861 GAAGTTGACCGGCTGGGAGGAGTACCGTCCGAAAGATTAAATCCCAAGGAATTGACGGGAC 920
Oy 899 CCGCAAGAGGAGGAGTATGATGATTAATTCGATGACGAGGAGGAGGAGGAGGAGGAGG 958
Db 921 CCGCAAGAGGAGGAGTATGATGATTAATTCGATGACGAGGAGGAGGAGGAGGAGGAGG 980
```

```
Oy 959 TTGACATGTAAGCGAATTTTCTAGAGATAGATTAGTC---TTGGGGAACGCTAACACAGG 1015
Db 981 TTGACATGCTTGGGAATCTTAAGAGACATTAAGAGTGTGCCGAAAGGAGCCCAAGACACAGG 1040
Oy 1016 TGCTGATGAGCTGTGCTGACGCTCGTGTGAGAGATGTTGGGTTAATGCTCCGCAACGAGCG 1075
Db 1041 TGCTGATGAGCTGTGCTGACGCTCGTGTGAGAGATGTTGGGTTAATGCTCCGCAACGAGCG 1100
Oy 1076 CAACCTTTGCTAATTAATTCGATCATTTT--GGTTGGGCACTTTAATGAGACTGCGGCTGAC 1134
Db 1101 CAACCTTTGCTAATTAATTCGATCATTTAATAATGAGCACTTTAGTGAAGACTGCGGCTGAC 1160
Oy 1135 AAACCGGAGGAAGGTGGGAGTGAAGTCAAGTCTCAATGAGCCCTTAATGGGTGGGCTTAC 1194
Db 1161 AAACCGGAGGAAGGTGGGAGTGAAGTCAAGTCTCAATGAGCCCTTAATGGGTGGGCTTAC 1220
Oy 1195 ACGTATATCAATGAGCGGCTACAGAGGAGGTTGGCAACCGCGAGGGGAGGAGTAACTCAGAA 1254
Db 1221 ACGTATATCAATGAGCGGCTACAGAGGAGGTTGGCAACCGCGAGGGGAGGAGTAACTCAGAA 1280
Oy 1255 AGCGGCTGTATGCTCGGATCGGAGTCTGCAACTGCACTCGGTGAAGTCCGAAATCGCTAAT 1314
Db 1281 AGCGGCTGTATGCTCGGATCGGAGTCTGCAACTGCACTCGGTGAAGTCCGAAATCGCTAAT 1340
Oy 1315 AATGCGGATACAGCATGTCGGGAGGAGTAAAGTTCGGGAGTCTGCAACACCGGCGCTCA 1374
Db 1341 AATGCGGATACAGCATGTCGGGAGGAGTAAAGTTCGGGAGTCTGCAACACCGGCGCTCA 1400
Oy 1375 CACCATGGGAGTGGGTTTCAACAGAGCAGGATGCTTAACCGTAAGAGAGGCGCTTGCCA 1434
Db 1401 CACCATGGGAGTGGGTTTCAACAGAGCAGGATGCTTAACCGCAAGAGAGGCGCTTGCCA 1460
Oy 1435 CGGTGAATTCATGATCGGGGTG 1457
Db 1461 CGGTGGGGGTGATGATCGGGGTG 1483
```

RESULT 2  
US-11-198-242-2  
; Sequence 2, Application US/11198242  
; Publication No. US2006003545A1  
; GENERAL INFORMATION:  
; APPLICANT: AJINOMOTO CO., INC.  
; TITLE OF INVENTION: Process for the production of beta-amino acids by using acylase  
; FILE REFERENCE: AB04037  
; CURRENT APPLICATION NUMBER: US/11/198, 242  
; CURRENT FILING DATE: 2005-08-08  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1522  
; TYPE: DNA  
; ORGANISM: Burkholderia sp.  
; FEATURE:  
; NAME/KEY: 16S rDNA  
; LOCATION: (1)..(1522)  
; OTHER INFORMATION:  
US-11-198-242-2

Query: March 78.9%; Score 1149.2; DB 18; Length 1522;  
Best Local Similarity 88.2%; Pred. No. 0; Mismatches 163; Indels 10; Gaps 3;  
Matches 1287; Conservative 0;

```
Oy 1 ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGCAGGATCTTGAT 60
Db 20 ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGGAGGAGGAGGAGGAGG 77
Oy 61 CTGTGGCGAGTGGCGGACGGGTGAGTAATGATCGGAACGTATCCAGAAAGGGGGGTA 120
Db 78 CTGGCGGCGAGTGGCGGAACGGGTGAGTAATGATCGGAACGTATCCGATGAGTGGGGGATA 137
Oy 121 ACGCATGGAAGATGCTTAATCCGCATATATCTTAAGAGGAGGAACGAGGAGATCGAAA 180
```





421 AAAAGTTACGGTAATTAATCGTACTGACATGAGGTAATGACAGAAAGACCGGCTAAC 480  
438 AACGGCTCTTCTTAATAAAGAGGGCTAAATGACGTAACGTAAGATTAAGCACCGGCTAAC 497  
481 TACGTCACAGACGCGCGGTAAATAGTAGAGGTGCAAGCTTAATCGAATTACTGGGCGT 540  
488 TACGTCACAGACGCGCGGTAAATAGTAGAGGTGCAAGCTTAATCGAATTACTGGGCGT 557  
541 AAAAGGTGCGACGCGCGCTTTGTAAATCAGATGTGAATCCCGGGCTTAACCTGGGAAT 600  
558 AAAAGGTGCGACGCGCGGTAAAGACAGTTGTGAATCCCGGGCTTAACCTGGGAAC 617  
601 TGGCTTTGAAACTAAGGCTAGAGTGTGGGCGAGAGGAGTGAATTCATGATGTAAGAG 660  
618 TGCATCTGAGTGAATCGCTGAGATGACGAGAGGGAGTGAATTCGCGGTGAGCAG 677  
661 TGAATGCGTAGAGATATGGAAGAAATCATGATGCGAAGGCGACGCTCTGGGTTAACT 720  
678 TGAATGCGTAGATATGCGAGAGAAACCGATGCGAAGGCGAATCCCTGGGCGCTGTA 737  
721 GACGCTCATGACGAAAGCGTGGGAGACAAACAGATTAGATACCTGGTACTCAAGCC 780  
728 GACGCTCATGACGAAAGCGTGGGAGACAAACAGATTAGATACCTGGTACTCAAGCC 797  
781 CTAACGATGTCACATGTTGTGGGCGCTTATTAAGGCTGTGAAGAGTAAAGCGGTGA 840  
798 CTAACGATGTCACATGTTGTGGGCGCTTATTAAGGCTGTGAAGAGTAAAGCGGTGA 857  
841 AGTTGACCGCTGGGAGAGTACCGTGCAGAAATTAATACTCAAGAAATTTGACGGGAGCC 900  
858 AGTTGACCGCTGGGAGAGTACCGTGCAGAAATTTGAACTCAAGAAATTTGACGGGAGCC 917  
901 GCACAGAGGCGGTGATTAATGTGGATTAATTCGATGCAAGCGGAAATCTTACCTT 960  
918 GCACAGAGGCGGTGATTAATGTGGATTAATTCGATGCAAGCGGAAATCTTACCTT 977  
961 GACATGTAAGCGAATTTTCTAGAGATAGATTAAGTGTCT---TGGGAAAGCTAACAGAGGT 1017  
978 GACATGTAAGCGAATTTTCTAGAGATAGATTAAGTGTCTCGAAGAAAGCGTAAACAGAGGT 1037  
1018 CTGCAATGCTGTCTGCACTGCTGTCTGAGATGTTGGTTAAAGTCCGCAACGACGCA 1077  
1038 CTGCAATGCTGTCTGCACTGCTGTCTGAGATGTTGGTTAAAGTCCGCAACGACGCA 1097  
1078 ACCCTTGTCAATTAATGCGATATTTGGTGGGCACTTTAATGAATCTGCGGTGACAA 1137  
1098 ACCCTTGTCAATTAATGCGATATTTGGTGGGCACTTTAATGAATCTGCGGTGACAA 1156  
1138 CCGAGAGAGGTGGGAGATGAGCTCAAGTCTCATAGGCCCTTAATGGGTAGGGCTTACAG 1197  
1157 CCGAGAGAGGTGGGAGATGAGCTCAAGTCTCATAGGCCCTTAATGGGTAGGGCTTACAG 1216  
1198 TAAATCAATGCGCGCTACAGAGGTTGCGAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1257  
1217 TCATCAATATGCTGTGTACAAAGGTTGCGAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1276  
1258 GCGTGTATGCTGGATGCGAGTCTGCAACTGCACTTCGTAAGTGGAAATGCTGTAT 1317  
1277 CAGTGTATGCTGGATGCGAGTCTGCAACTGCACTTCGTAAGTGGAAATGCTGTAT 1336  
1318 CCGGATACAGATGTCGCGGTGAATACGTTCCCGGCTTGTGACACCGCCCGCTACAC 1377  
1337 CCGGATACAGATGTCGCGGTGAATACGTTCCCGGCTTGTGACACCGCCCGCTACAC 1396  
1378 CATGGAGTGGGTTTCAACAGAAAGGATGTAATCAACGTAAGAGAGGCGCTTGCACGG 1437  
1397 CATGGAGTGGGTTTCAACAGAAAGGATGTAATCAACGTAAGAGAGGCGATTAACAGG 1456  
1438 TGAATTCATGACTGGGCTG 1457  
1457 CAGGTTCTGTGACTGGGCTG 1476

RESULT 4  
US-10-515-311-5  
; Sequence 5, Application US/10515311  
; Publication No. US2006030021A1  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Susan J  
; TITLE OF INVENTION: Novel Microorganisms and Uses Therefor  
; FILE REFERENCE: 11752-006US1  
; CURRENT APPLICATION NUMBER: US/10/515,311  
; PRIOR FILING DATE: 2004-11-22  
; PRIOR APPLICATION NUMBER: PCT/NZ2003/000100  
; PRIOR FILING DATE: 2003-05-22  
; PRIOR APPLICATION NUMBER: NZ 519137  
; PRIOR FILING DATE: 2002-05-22  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1405  
; TYPE: DNA  
; ORGANISM: Acidithiobacillus  
US-10-515-311-5  
Query Match 74.8%; Score 1089.6; DB 11; Length 1405;  
Best Local Similarity 87.2%; Pred. No. 0;  
Matches 1233; Conservative 0; Mismatches 169; Indels 12; Gaps 3;  
18 TGCCTTACATGCAAGTCCGAGCGGACGAGATGCTTGATCTGTTGGGAGTGGCGG 77  
1 TGCCTTACATGCAAGTCCGAGCGGACGAGATGCTTGATCTGTTGGGAGTGGCGG 56  
78 ACGGAGTGAATGATCATGCGAAGCTATCAGAAAGAGGGGGTAAAGCATCGAAGATGTG 137  
57 ACGGAGTGAATGATCATGCGAAGCTATCAGAAAGAGGGGGTAAAGCATCGAAGATGTG 116  
138 CTAATACCGCATATCTCTTAAGAGAAAGAGGAGGAGTCAAGAACCTTGGCTTTTGA 197  
117 CTAATACCGCATATCTCTTAAGAGAGTGAAGAGGAGGAGTCAAGAACCTTGGCTTTTGA 176  
198 GCGGCGCATGCTGATTAATGCTGATGTTGGTGGTAAAGGCTTAACGAGCGGATGAGTA 257  
177 GCGGCGCATGCTGATTAATGCTGATGTTGGTGGTAAAGGCTTAACGAGCGGATGAGTA 236  
258 GTTGATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 317  
237 GCTGATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 296  
318 AGGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 377  
297 AGGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 356  
378 TGAAGAGGCGCTTGGGCTGTTGAAGCTCTTCACTCGAAGAAAGGTTAAGGTTAATA 437  
357 GGAATGAAGGCGCTTGGGCTGTTGAAGCTCTTCACTCGAAGAAAGGTTAAGGTTAATA 416  
438 ATCTGATCTGAGCGGATGAGCAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 497  
417 CTTGGGCTCATGAGCGGATGAGCAGAAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 476  
498 GGTATATGTAAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 557  
477 GGTATATGTAAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 536  
558 CTTTGTATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 617  
537 TTAATATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 596  
618 GGTATATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 677  
597 AGCTAGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 656  
678 TGAAGAAATCATGATGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 737  
657 CGAGAGAAACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 716





```
Db      919  GCACAGGCGGTGGAGCATGTGTTTAATTCGAGAGACGCCAGAACCTTACAGAGCCT 978
Qy      960  TACACATGTAGCGAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTTAACACAGGTGC 1018
Db      979  TACACATGTAGAGAACCTTTCAGAGATGATGTGTGCTTCGGGAACTCTGACACAGGTGC 1038
Qy     1019  TGCATGGCTGTGTGCTGAGCTGTGTGTGATGTTAGTGGGTTAAGTCCCGCAACAGAGCGCA 1078
Db     1039  TGCATGGCTGTGTGCTGAGCTGTGTGTGATGTTAGTGGGTTAAGTCCCGCAACAGAGCGCA 1098
Qy     1079  CCCTTGTCAATTAATTCGCATCA--TTTGGTTGGGCACTTAAATGAGACTGCCGCTGACAA 1136
Db     1099  CCCTTGTCTTAGTTAGTACAGACGATATGTGGGCACTTAAGAGAGCTGCCGCTGACAA 1158
Qy     1137  ACCGAGAGAGGTGGGGATGAGCTCAAGTCTTCATAGCCCTTAATGGGTAAGGCTTCAAC 1196
Db     1159  ACCGAGAGAGGTGGGGATGAGCTCAAGTCAATCATAGCCCTTAATGGGCTTCAACAC 1218
Qy     1197  GTAATACATAGCGGTGACAGAGGTTGCCAACCGCGAGGGGAGCTTAATCTCAGAAAG 1256
Db     1219  GTGCTACAAATGGTCCGTACAGAGGTTGCCAAGCCGCGAGGTGAGCTAATCTCACAAAA 1278
Qy     1257  CGCGTGTAGTCCGAGTCGAGTCTGCAACTCGACTCCGTGAAGTCGGAATCGTAGTAA 1316
Db     1279  CGGATCGTAGTCCGAGTCGAGTCTGCAACTCGACTCGGTGAAGTCGGAATCGTAGTAA 1338
Qy     1317  TCGCGATACGATGTCCGCGGTGAATACGTTCCCGGCTTTTGTACACACCGCCGCTACA 1376
Db     1339  TCGCAATACGAAATGTCCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCGCTACA 1398
Qy     1377  CCATGGAGATGGGTTTACACAGAGAGGTAGTCTTAACCGTAAGAGGCGCTTGCCACG 1436
Db     1399  CCATGGAGATGGGTTTGCACACAGAGGTAGTCTTAACCTTGGAAGAGCGGTTACACAG 1458
Qy      1437  GTGAGATTCATGACTGGGGTG 1457
Db      1459  GTGTGATTCATGACTGGGGTG 1479

RESULT 6
US-10-831-286A-48683
; Sequence 48683, Application US/10831286A
; Publication No. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831.286A
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; NUMBER OF SEQ ID NOS: 2003-04-24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48683
; LENGTH: 1493
; TYPE: DNA
; ORGANISM: Pantoea dispersa
US-10-831-286A-48683

Query Match      70.1%; Score 1021.2; DB 11; Length 1493;
Best Local Similarity 82.9%; Pred. No. 5,2e-309;
Matches 1214; Conservative 0; Mismatches 243; Indels 7; Gaps 4;
```

```
Db      61  TCTTTGGGTGGCGATGGCGGACCGGTGAGTAATGTCTGGGAAACTGCCCATGATGAGGG 120
Qy     117  GGTAAACGATGAAAGATGTCTTAATACCGCATATATCTTAAGAGAAAGACAGGGATC 176
Db     121  GATTAATCTAGAAAACGGTACTAATACCGCATACGTCGCAACCAAAAGTGGGGAGCC 180
Qy     177  GAAAGACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTGGGAGTAAAGGC 236
Db     181  TTCGGGCTTCAACACCATGGATGTGCCCCAGATGGGAAATTAGCTAATAGTGGGTAATG 240
Qy     237  CTACCAAGCCGACGATCAGTAGTGTGCTGAGAGAGACCAAGCACACTGGGACTGAGA 296
Db     241  TCACCTAGGCGACATCTCCCTAGCTGTGTGTGAGAGATGACCAAGCACACTGAACTG 300
Qy     297  CACGGCCCAACTCTTACCGGAGGACGATGAGGGAATTTTGAACAATGGGCGAAGCT 356
Db     301  CACGCTCCAGACTCTTACCGGAGGACGATGAGGGAATTTGCAAAATGGGCGCAAGCT 360
Qy     357  GATCCAGCAATGCCGCGTGAAGTGAAGAGGCTTCGGGTTGTAAGCTCTTCACTGAG 416
Db     361  GATGACGCAATGCCGCGTGAAGTGAAGAGGCTTCGGGTTGTAAGCTCTTCACTGAG 420
Qy     417  AAGAAAGGTTACCGTAAATATCGTACATGACGATATCGACAGAAAGACACCGGC 476
Db     421  AGAAAGGCGGTGAGGTTAATTAACCTTGCCGATTTGACGTTACCGCAGAAAGACAC 480
Qy     477  TAACTAGTGTGCAGACAGCCGCGGTAAATACGTAAGGTGCAAGGCTTAATCGGAAT 536
Db     481  TAACTCGGTGCACAGCCGCGGTAAATACGTAAGGTGCAAGGCTTAATCGGAAT 540
Qy     537  GCGTAAGGCGGCGCAGGCGGCTTGTAAAGTCAATGTAAGTATCCCGGGCTTAACCT 596
Db     541  GCGTAAGGCGCAGGCGGCTTGTAAAGTCAATGTAAGTATCCCGGGCTTAACCT 600
Qy     597  GAATTCGTTGAAACTTAACAAGCTAGATGTGACAGAGGAGGTGAAATTCATGTGTA 656
Db     601  GAACTGATTTGAAACTGTGAGAGCTTAGATCTGTAGAGGGGGGTAGAAATTCAGGTG 660
Qy     657  GCACTGAATTCGTAGATATGGAAGAACATGCAATGCGGACGACCTCTCGGTTAA 716
Db     661  GCGGTGAATTCGTAGATATGGAAGAACATGCAATGCGGACGACCTCTCGGTTAA 720
Qy     717  CACTGACGCTCATGACAGAAAGCGTGGGAGCAAGATTAGTACCCTGATGATGCA 776
Db     721  GACTGACGCTCATGATGGAAGCGTGGGAGCAAGATTAGTACCCTGATGATGCA 780
Qy     777  CGCCCTAAACGATGCTCAATGTTGTGGGCTTATTAGGCTTGG-7AACGAAGCTAACG 835
Db     781  CGCCCTAAACGATGCTCAATGTTGTGGGCTTATTAGGCTTGG-7AACGAAGCTAACG 840
Qy     836  CGTGAAGTTGACCGCTCGGGAGTACGCGTGCAGAAATTAAACTCAAGAAATTGACGG 895
Db     841  CGTGAAGTTGACCGCTCGGGAGTACGCGTGCAGAAATTAAACTCAATGAATTGACGG 900
Qy     896  GACCCGCAAGCGGTGATTAATGATTAATTCGATGGAACCGGAAAAACCTTACTTA 955
Db     901  GACCCGCAAGCGGTGATTAATGATTAATTCGATGGAACCGGAAAAACCTTACTTA 960
Qy     956  CCCTTGACATGATGCAATTTTCTAGAGATGATTAAGTGTG-CTTCGGGAACGCTAACAG 1014
Db     961  GCCTTGACATGCAAGAACTTAGCAGAGATGCTTTGTGCTTCGGGAACTCTGAGACAG 1020
Qy     1015  GTGCTGATGAGTGTGTGCTGAGCTGTGTGTGATGATTTGGGTTAAGTCCCGCAACAG 1074
Db     1021  GTGCTGATGAGTGTGTGCTGAGCTGTGTGTGAAATTTGGGTTAAGTCCCGCAACAG 1080
Qy     1075  GCAACCTTGTCAATTAATTCGCAATC-ATTGTGTTGGCACTTTAAATGAACTGCGG 1133
Db     1081  GCAACCTTGTCAATTAATTCGCAATC-ATTGTGTTGGCACTTTAAATGAACTGCGG 1140
Qy     1134  CAAACCGGAAGAGGTGGGATGACGTCAGTCTCTCAATGCGCTTATGGGTTAGGGCTTCA 1193
```





```
FILE REFERENCE: 032796-174.001
CURRENT APPLICATION NUMBER: US/10/831.286A
CURRENT FILING DATE: 2004-04-26
PRIOR APPLICATION NUMBER: 60/464,955
PRIOR FILING DATE: 2003-04-24
NUMBER OF SEQ ID NOS: 48788
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 48681
LENGTH: 1494
TYPE: DNA
ORGANISM: Pantoea dispersa
US-10-831-286A-48681

Query Match      70.0%; Score 1019.6; DB 11; Length 1494;
Best Local Similarity 82.9%; Pred. No. 1.6e-308;
Matches 1213; Conservative 0; Mismatches 244; Indels 7; Gaps 4;

QY 1 ATTTGAACGCTGGCGCGCATCTTTTACATGCAATGCAAGCGGACGACGAGATGCTTG-- 57
DB 1 ATTTGAACGCTGGCGCGCATCTTTTACATGCAATGCAAGCGGACGACGAGATGCTTG 60
QY 58 -CATCTGGTGGCGAGTGGCGGACGGGTGATGATGCAAGCATCGAAGAGCTTG 116
DB 61 TCTTTGGGTGGCGAGTGGCGGACGGGTGATGATGCTGGGAAATGCTCCGATGAGGG 120
QY 117 GGTAAACGATCGAAGATGCTTAATACCGCATATTAATCTTAAGAGGAAAGAGGGATC 176
DB 121 GATTAATCTGAAACGATGCTTAATACCGCATATTAATCTTAAGAGGAAAGAGGGATC 180
QY 177 GAAAGCCTTGGCGCTTTTGAAGCGGCGATGCTGATTAGCTAATGTTGGGTAAAGCC 236
DB 181 TTGGGCGCTCACACCATCGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 240
QY 237 CTACACGAGCGGACGATCATATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 296
DB 241 TCACCTAAGCGGACGATCCTAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 300
QY 297 CACGCGCCGAGACTCTTCTAAGCGGAGGACGAGTGGGGAAATTTTGAACATGGGCGCAAGCT 356
DB 301 CACGCGCCGAGACTCTTCTAAGCGGAGGACGAGTGGGGAAATTTTGAACATGGGCGCAAGCT 360
QY 357 GATTCAGCAATGCGCGCTGATGAGAGAGGCGCTTGGGTTTGAAGCTTTTCAATCGAG 416
DB 361 GATTCAGCAATGCGCGCTGATGAGAGAGGCGCTTGGGTTTGAAGCTTTTCAATCGAG 420
QY 417 AAGAAAGCTTACGCTAATTAATCGTCACTGATGAGGCTATGACAGAAAGACCGGC 476
DB 421 AAGAAAGCTTACGCTAATTAATCGTCACTGATGAGGCTATGACAGAAAGACCGGC 480
QY 477 TAATCTAGTGGCGAGCGCGGTAATAGTAGAGGTGCAAGCTTATCGAATTTACTG 536
DB 481 TAATCTAGTGGCGAGCGCGGTAATAGTAGAGGTGCAAGCTTATCGAATTTACTG 540
QY 537 GCGTAAAGGGTGGCGAGCGGCTTTTGAAGTGAATGGAATCCCGGCGCTTAACCTGG 596
DB 541 GCGTAAAGGGTGGCGAGCGGCTTTTGAAGTGAATGGAATCCCGGCGCTTAACCTGG 600
QY 597 GAATTCGCTTTGAATCTAAGGCTAGAGTGGCGAGAGAGAGTGAATTCATGATGTA 656
DB 601 GAATTCGCTTTGAATCTAAGGCTAGAGTGGCGAGAGAGTGAATTCATGATGTA 660
QY 657 GCAGTAAATGCTAGAGATATGGAAGAAATCATGATGGCGAAGGAGCGCTCTGGGTAA 716
DB 661 GCAGTAAATGCTAGAGATATGGAAGAAATCATGATGGCGAAGGAGCGCTCTGGGTAA 720
QY 717 CACTGACGCTCATGCAAGAAAGCGTGGAGAGCAAAAGATTAATGATACCTGTGATGCA 776
DB 721 GACTGACGCTCATGAGTGAAGAGTGGAGAAATCCCGTGGTGGTGGTGGTGGTGGTGG 780
QY 777 CGCCCTTAAGAGATGCAATGATTTGGTGGCTTTATTAAGCTTGG--TAAGAGAGCTAAG 835
DB 781 CGCCCTTAAGAGATGCAATGATTTGGTGGCTTTATTAAGCTTGGCTTGGAGCTTAAG 840
```

```
QY 836 CGTGAAGTTGACCGCTGGGAGATGACGTCAGCAAGATTAAACTCAAGAAATTGACGG 895
DB 841 CGTTAAGTCGACCGCTGGGAGATGACGTCAGCAAGATTAAACTCAAGAAATTGACGG 900
QY 896 GACCCGACAAAGCGGTGATTTATGATTTAATTCATGCAAGCGGAAATCTTACTTA 955
DB 901 GACCCGACAAAGCGGTGATTTATGATTTAATTCATGCAAGCGGAAATCTTACTTA 960
QY 956 CCTTTGACATGATGCAATTTTCTAGAGATGATTAAGT--CTTGGGAAACCTTAACAG 1014
DB 961 GCTTTGACATGATGCAATTTTCTAGAGATGATTAAGT--CTTGGGAAACCTTAACAG 1020
QY 1015 GTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074
DB 1021 GTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1075 GCAACCTTGTCAATTAATTTGCAATC-ATTGCTTGGGCACTTTATGAGACTGCGGTGA 1133
DB 1081 GCAACCTTGTCAATTTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1134 CAACCGGAGGAGGTGGGATGACGTCAGTCTCATGCGCTTATGAGGCTTCA 1193
DB 1141 TAACCGGAGGAGGTGGGATGACGTCAGTCTCATGCGCTTATGAGGCTTCA 1200
QY 1194 CACGTAATCAATGGCGGTGACAGAGGTTGCCAACCGGAGGGGAGCTTAATCTAGA 1253
DB 1201 CACGTAATCAATGGCGGTGACAGAGGTTGCCAACCGGAGGGGAGCTTAATCTAGA 1260
QY 1254 AAGCGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1313
DB 1261 AAGCGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1314 TAATGCGGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1373
DB 1321 TAATGCGGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
QY 1374 ACACATGAGGATGGGTTTACACAGAGGTAATCTTAACCGTAAAGGAGGCGCTTGGC 1433
DB 1381 ACACATGAGGATGGGTTTACACAGAGGTAATCTTAACCGTAAAGGAGGCGCTTACC 1440
QY 1434 ACGGTGATGATGATGCTGCGGTG 1457
DB 1441 ACTTTGATGATGATGATGCTGCGGTG 1464
```

```
RESULT 10
US-10-831-286A-48677
Sequence 48677, Application US/10831286A
Publication No. US20060046246A1
GENERAL INFORMATION:
APPLICANT: ZENG, QIANGDONG
APPLICANT: CHATELIER, SONIA
APPLICANT: MOIR, DONALD T.
APPLICANT: LACROIX, BRUNA
APPLICANT: CHILDRESS, DARRELL
TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
FILE REFERENCE: 032796-174.001
CURRENT APPLICATION NUMBER: US/10/831.286A
CURRENT FILING DATE: 2004-04-26
PRIOR APPLICATION NUMBER: 60/464,955
NUMBER OF SEQ ID NOS: 48788
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 48677
LENGTH: 1507
TYPE: DNA
ORGANISM: Providencia rettgeri
US-10-831-286A-48677
```

```
Query Match      69.9%; Score 1018; DB 11; Length 1507;
Best Local Similarity 83.2%; Pred. No. 5.2e-308;
Matches 1217; Conservative 0; Mismatches 240; Indels 5; Gaps 5;
```

1 ATTGAACGCTGGCGGCGATCTTTACACATGCAAGTCGAACCGGACGAC -GGATGCTTGCA 59  
11 ATTGAACGCTGGCGGCGGCGCTTAACACATGCAAGTCGAACCGGCTTAACGAGGAACTTGCT 70  
60 TCTGG -TGGCGAGTGGCGGACGCGGTGAATGCAATCGGAACGTAATCCAGAAAGGAGG 118  
71 TCTGCTGACGAGCGGCGGACGCGGTGAATGTAATGGGATCTGCGCCGATGAGAGGGGGA 130  
119 TAAACGATGAAAGATGCTTAATACCGCATATCTTAAAGAGGAAACGAGGATGCA 178  
131 TAAACATGGAAGACGCTGCTTAATCCGATATCTTAAAGAGGAAACGAGGAACTT 190  
179 AAGACCTTGCGCTTTTGAGCGCGCATGTCTGATTAAGTGAATGAGTGAAGGCTT 238  
191 CGGTCTTGCGCTTAATGCAATGCAATGAGATTAAGTGAATGAGTGAAGGCTT 250  
239 ACCAAGGCGACGATCAGTATGCTGAGAGGACGACCGACACCTGGAAGTGAACA 298  
251 ACTGAGGCGACGATCCCTGAGTGTGAGAGGATGATGACGACACCTGGAAGTGAACA 310  
299 CGGCGGACGCTCTTACGAGGAGGACGAGTGGGGAATTTTGACATGAGCGCAAGCTGA 358  
311 CGGCGGACGCTCTTACGAGGAGGACGAGTGGGGAATTTTGACATGAGCGCAAGCTGA 370  
359 TCCAGCAATGCGCGGAGTGAAGAGGCTTCCGGTTGTAAGGCTCTTCACTGAGGAA 418  
371 TCCAGCAATGCGCGGAGTGAAGAGGCTTCCGGTTGTAAGGCTCTTCACTGAGGAA 430  
419 GAAAGGTTACGCTTAATATGCTGATCTGATGACGCTATGACAGAGAAAGCAGCGCTA 478  
431 GAAAGGTTGATGCTTAATATGCTGATCTGATGACGCTATGACAGAGAAAGCAGCGCTA 490  
479 ACTGAGTGGCAGACGCGCGGTGAATGCTGAGGCTGACAGGCTTAATCGAAATTAATCTGG 538  
491 ACTGAGTGGCAGACGCGCGGTGAATGCTGAGGCTGACAGGCTTAATCGAAATTAATCTGG 550  
539 GTPAAGGAGGCGGCGGCTTTGTAAGTGAATGCAATGCAATGCAATGCAATGCAATGCAAT 598  
551 GTPAAGGAGGCGGCGGCTTTGTAAGTGAATGCAATGCAATGCAATGCAATGCAATGCAAT 610  
599 ATTGCGTTGAAACTTACAGAGCTGAGTGGCAGAGGAGGAGTGAATTCATGCTGAGC 658  
611 ATGGCATCTAAGACGCTGAGTGAAGTCTTGAAGAGGAGGAGTGAATTCATGCTGAGC 670  
659 AGTGAATGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 718  
671 GGTGAATGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 730  
719 CTGAGGCTCATGACGAAAGCGTGGGAGCAACAGGATTAATGATCCCTGATGCTGACG 778  
731 CTGAGGCTCATGACGAAAGCGTGGGAGCAACAGGATTAATGATCCCTGATGCTGACG 790  
779 CCTTAAACGATGTC -AACTAGTGTGTTGGCTTATTAAGCTTGTGTAACGAACTTAACGCG 837  
791 CTGTAACGATGTCGATTTTGAAGTGTTCCTTAAGAGGATGCTTCCGGAACCTTAACGCG 850  
838 TGAAGTGAACGCGCTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 897  
851 TTAATTCGACCGCTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 910  
898 CCGGACAAAGGAGTGAATGATGATTAATTCATGCAACGCGAAACCTTAACCTTACC 957  
911 CCGGACAAAGGAGTGAATGATGATTAATTCATGCAACGCGAAACCTTAACCTTACC 970  
958 CTGCAATGTAAGCAATTTTCTAGAGATGATTAAGT -CTTGGGAGCGCTTAACAGAGT 1016  
971 CTGCAATGTAAGCAATTTTCTAGAGATGATTAAGT -CTTGGGAGCGCTTAACAGAGT 1030  
1017 GCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076  
1031 GCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090  
1077 AACCTTGTCAATTAATGCTATC -ATTGCTGTTGGGACCTTAAATGAGACTGCGGTTGACA 1135

1091 AACCTTATCTTTGTTGCCAGCATACGCTGGGAACTCAAGAGACTGCGGTGATA 1150  
1136 AACCGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1195  
1151 AACCGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1210  
1196 GGTATATCAATGAGGCGGACGAGAGGTTGCCAACCGCGAGAGGAGGAGTGAATCTCAGAAA 1255  
1211 GGTGCTACATGAGGCTATCAAAAGAGGACCTCCGAGAGCAAGCGAACTCATTA 1270  
1256 GCGGCTGATGCTCGGATGCGAGTCTGCAACTGCACTCGGAAAGTGGAAATGCTGATGA 1315  
1271 GTAGGTGATGCTCGGATGCGAGTCTGCAACTGCACTCGGAAAGTGGAAATGCTGATGA 1330  
1316 ATCCGATGACGATGCTGCGGTGAATGCTTCCCGGCTTGTGACACGCGCGCTGAC 1375  
1331 ATCCGATGACGATGCTGCGGTGAATGCTTCCCGGCTTGTGACACGCGCGCTGAC 1390  
1376 ACCATGGAGTGGGTTTCAACAGAGGATGCTAACCGTAAAGAGGCGCTTGCCAC 1435  
1391 ACCATGGAGTGGGTTTCAACAGAGGATGCTAACCGTAAAGAGGCGCTTGCCAC 1450  
1436 GGTGATTCATGACTGGGCTG 1457  
1451 TTTGATTCATGACTGGGCTG 1472  
RESULT 11  
US-10-831-286A-48684  
; Sequence 48684, Application US/10831286A  
; Publication No. US2006046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILDRESS, DARRELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; CURRENT FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; PRIOR FILING DATE: 2003-04-24  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 48684  
; LENGTH: 1508  
; TYPE: DNA  
; ORGANISM: Providencia rustigianii  
US-10-831-286A-48684  
Query Match 69.8%; Score 1017; DB 11; Length 1508;  
Best Local Similarity 83.2%; Pred. No. 1.1e-307;  
Matches 1217; Conservative 0; Mismatches 240; Indels 6; Gaps 5;  
1 ATTGAACGCTGGCGGCGATCTTTACACATGCAAGTCGAACCGGACGAC -GGATGCTTGCA 59  
11 ATTGAACGCTGGCGGCGGCGCTTAACACATGCAAGTCGAACCGGCTTAACGAGGAACTTGCT 70  
60 TCTGG -TGGCGAGTGGCGGACGCGGTGAATGCAATCGGAACGTAATCCAGAAAGGAGG 118  
71 TCTGCTGACGAGCGGCGGACGCGGTGAATGTAATGGGATCTGCGCCGATGAGAGGGGGA 130  
119 TAAACGATGAAAGATGCTTAATACCGCATATCTTAAAGAGGAAACGAGGATGCA 178  
131 TAAACATGGAAGACGCTGCTTAATCCGATATCTTAAAGAGGAAACGAGGAACTT 190  
179 AAGACCTTGCGCTTTTGAGCGCGCATGTCTGATTAAGTGAATGAGTGAAGGCTT 238  
191 CGGTCTTGCGCTTAATGCAATGCAATGAGATTAAGTGAATGAGTGAAGGCTT 250  
239 ACCAAGGCGACGATCAGTATGCTGAGAGGACGACCGACACCTGGAAGTGAACA 298

Db 251 ACCAAGGGGAGGATCCCTAGCTGCTCTAGAGAGATGATCAGCACTGGGACTGAGCA 310  
 Qy 259 CGGGCCAGACTCTCTAGGGAGGAGCAGATGGGGAAATTTTGGACAATGGGCGCAACCTTGA 358  
 Db 311 CGGGCCAGACTCTCTAGGGAGGAGCAGATGGGGAAATTTTGGACAATGGGCGCAACCTTGA 370  
 Qy 359 TCCAGCAATGCGCGCTGAGTGAAGAAAGGCTTCGGGTTGTAAAGCTCTTTTCAGTCAAGAA 418  
 Db 371 TGCACCAATGCGCGCTGAGTGAAGAAAGGCTTCAGGTTGTAAAGTACTTTTCAGTTGGAG 430  
 Qy 419 GAAAAGTTACGCTTAATATATCTGATCTATGACGCTATCCAGAAAGAGACCGGCTA 478  
 Db 431 GAAGGCGTGTGATGCTAATATCTATCAGCCGATTAACCTTACCAAGAAAGAGACCGGCTA 490  
 Qy 479 ACTAGGTGCAAGCAGCGCGGTATATACGAGGGGTGCAAGCGTAAATCCGAAATTAATGAGG 538  
 Db 491 ACTCGGTGCAAGCAGCGCGGTATATACGAGGGGTGCAAGCGTAAATCCGAAATTAATGAGG 550  
 Qy 539 GTAAAGGTGCGCAGCGCGCTTTGTAGTCAATGATGAAATCCCGGCGCTTAACCTGAGA 598  
 Db 551 GTAAAGCGCAGCGCGCTTTGTAGTCAATGATGAAATCCCGGCGCTTAACCTGAGA 610  
 Qy 599 ATTGCGTTGAAACTTAAAGCTGATGCTGAGGAGGAGGAGGAGTTCATGTTGATG 658  
 Db 611 ATGGCATCTAAGACTGTGCTAGAGTCTTGTAGAGGGGGGTAGAAATTCATGTTGATG 670  
 Qy 659 ACTGAAATCCGTAGAGATATGAGAAACATCATGAGCGAAAGGAGCGCTCTGAGTTAA 718  
 Db 671 GGTTAAATCCGTAGAGATATGAGAAATCCCGTGGCAAGGCGCGCTCTGAGTTAA 730  
 Qy 719 CTGAGCGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAAGTAACTCTGTAGTCCAG 778  
 Db 731 CTGAGCGCTCATGTCGAAAGCGTGGGAGCAAAACAGATTAAGTAACTCTGTAGTCCAG 790  
 Qy 779 CCTTAAACGATGTCAAATGTTGTTGGGCTTATTAAGCTTGG-1TACGAAAGTAAAGCG 837  
 Db 791 CTGTAAACGATGTCAAATGTTGTTGGGCTTATTAAGCTTGG-1TACGAAAGTAAAGCG 850  
 Qy 838 TGAAGCTGACCGCTTGGGAGTACCGGTCCCAAGATTAATACTCAAGAAATGAGCGGG 897  
 Db 851 TTAAATCGACCGCTTGGGAGTACCGGTCCCAAGATTAATACTCAAGAAATGAGCGGG 910  
 Qy 898 CCGGCAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAATACTTACCTTACC 957  
 Db 911 CCGGCAAGCGGTGATGATGATTAATTCGATGCAACGCGAATACTTACCTTACC 970  
 Qy 958 CTGACATGTAGCGAAATTTTCTAGAGATGATTAAGT-CTTGGGAAACGCTTACACAGT 1016  
 Db 971 CTGACATGTAGCGAAATTTTCTAGAGATGATTAAGT-CTTGGGAAACGCTTACACAGT 1030  
 Qy 1017 GCTGATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1076  
 Db 1031 GCTGATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1090  
 Qy 1077 AACCTTGTCAATTAATTTGCAATCA--TTTGGTTGGGCACTTTAATGAGACTGCGGTGAC 1134  
 Db 1091 AACCTTGTCAATTTTGTGTCAGACGCTATGATGAGAACTCAAAAGAGACTGCGGTGAT 1150  
 Qy 1135 AAACCGGAGAGAGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGGGAGTTCAC 1194  
 Db 1151 AAACCGGAGAGAGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGGGAGTTCAC 1210  
 Qy 1195 AGCTAATCAATGCGCGCTGATCAAGAGGTTCCCAACCGCGAGGGGAGCTAATCTAGAA 1254  
 Db 1211 ACCTGTCAATGCGCGCTGATCAAGAGGTTCCCAACCGCGAGGGGAGCTAATCTAGAA 1270  
 Qy 1255 AGCGCGTGTAGTCCGGAATGAGAGTCTGCAACTGCACTCCGTAAGTCCGGAATCGTACT 1314  
 Db 1271 AGTACGTGTAGTCCGGAATGAGAGTCTGCAACTGCACTCCGTAAGTCCGGAATCGTACT 1330  
 Qy 1315 AATCGCGATCAGCATGTGCGCGGTAAATACGTTCCCGGCTCTTGTACACACCGCGCTCA 1374

Db 1331 AATGTGATCATGAATGCTACGGTGAATACGTTCCCGGCGCTTGTACACACCGCCCTCA 1390  
 Qy 1375 CACCATGGAGTGGGTTTACACAGAACGATTAATCTTAACGTTAAGAGGGCGTTGCA 1434  
 Db 1391 CACCATGGAGTGGGTTTACACAGAACGATTAATCTTAACGTTAAGAGGGCGTTGCA 1450  
 Qy 1435 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1457  
 Db 1451 CTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1473  
 RESULT 12  
 US-10-831-286A-48682  
 ; Sequence 48682, Application US/10831286A  
 ; Publication No. US20060046246A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZENG, QIANDONG  
 ; APPLICANT: CHATELIER, SONIA  
 ; APPLICANT: MOIR, DONALD T.  
 ; APPLICANT: LACROIX, BRUNA  
 ; APPLICANT: CHILDRESS, DARRELL  
 ; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
 ; FILE REFERENCE: 032796-174.001  
 ; CURRENT APPLICATION NUMBER: US/10/831.286A  
 ; PRIOR FILING DATE: 2004-04-26  
 ; NUMBER OF SEQ ID NOS: 48788  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 48682  
 ; LENGTH: 1510  
 ; TYPE: DNA  
 ; ORGANISM: Moellerella wisconsensis  
 US-10-831-286A-48682  
 Query Match 69.7%; Score 1015.4; DB 11; Length 1510;  
 Best Local Similarity 82.7%; Pred. No. 3.4e-307;  
 Matches 1211; Conservative 0; Mismatches 246; Indels 8; Gaps 4;  
 Qy 1 ATTTGAACGCTGGCGGATGCTTTTACATGCAATGCAAGTGAACGGCAGCAGATGCTTGA- 59  
 Db 11 ATTTGAACGCTGGCGGATGCTTTTACATGCAATGCAAGTGAACGGCAGCAGATGCTTGA 70  
 Qy 60 ---TTGGTGGCGAGTGGCGGAGCGGCTGATGATTCGAACTTACGAAAGGG 116  
 Db 71 TTCTTTGCTGACGACCGGCGGAGCGGCTGATGATTCGAAAGGGATCTGCTGACAGAGGG 130  
 Qy 117 GGTAAACGATGGAAGATGCTTAATACCGCATATCTTAAGAGAGAAAGCAGGGATC 176  
 Db 131 GATTAATCTGGAAGACGTTAATACCGCATATCTTAAGAGAGAAAGCAGGGATC 190  
 Qy 177 GAAAGACCTTGGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGGTAAAGGC 236  
 Db 191 TTGGGCGCTTGGCGCTTGGAGTGAACCATATGGGATTAAGTGAAGGTAAAGGC 250  
 Qy 237 CTACCAAGCGCAGATCAGTATGTTGCTGAGAGAGACACAGCCACACTGGGACTGAGA 296  
 Db 251 TCACCTAGCGCAGATCTCTACCTGCTGAGAGAGATGATCAGCCACACTGAGACTGAGA 310  
 Qy 297 CACGCGCCAGATCTCTACCGGAGGAGCAGAGTGGGGAATTTTGAACATGGGCGCAAGCT 356  
 Db 311 CACGCGCCAGATCTCTACCGGAGGAGCAGAGTGGGGAATTTTGAACATGGGCGCAAGCT 370  
 Qy 357 GATTCAGCAATGCGCGTGTAGTGAAGAGAGCCCTTGGGTTGTAAAGCTTTCACTGAG 416  
 Db 371 GATTCAGCAATGCGCGTGTATGAAGAGAGCCCTTGGGTTGTAAAGCTTTCACTGAG 430  
 Qy 417 AAGAAAGGTTACGTTAATTAATCTGACTATGACGCTATGACAGAGAGAGAGAGAGAGAG 476  
 Db 431 AAGAAAGGCTTGTATTAATTAATCTGACTATGACGCTATGACAGAGAGAGAGAGAGAG 490  
 Qy 477 TAACTGCTGCGCAGAGCCGCGGTAAATAGTGGGTCGAAGGCTTAATCGAATTAATCTG 536



491 TAACTCCGTCGACGAGCCGCGGTAAATACGCGGGGTGCAAGCGTTAATCGAATTACTGG 550  
QY 537 GCGTAAAGGGTCCGACGCGCGCTTGTAACTGATGTAATCCCCGGGCTTAACTGG 596  
Db 551 GCGTAAAGGGACGACGCGCGCTTGTAACTGATGTAATCCCCGGGCTTAACTGG 610  
QY 597 GAATTGCGGTTAACTACAAAGGCTAGATGTTGGACGAGGGAGTGGAAATTCATGTTGA 656  
Db 611 GAATGCGATCTAAACTGCTACGCTAGACTCTTGTAGAGGGGGGTAGATTCATGTTGA 670  
QY 657 GCAAGTAAATGCGTAGATATGAAAGAACATCGATGGCGAAGGACGCTCTGGGTTAA 716  
Db 671 GCGGTAAATGCGTAGATATGAAAGAACATCGATGGCGAAGGACGCTCTGGGTTAA 730  
QY 717 CACTGACGCTCATGCAAGAACGTTGGGAGCAAAACGATTTAGTATCCCTGTGTTGCA 776  
Db 731 GACTGACGCTCATGCGGACGAAACGTTGGGAGCAAAACGATTTAGTATCCCTGTGTTGCA 790  
QY 777 GCGCCCTAAACGATGCACTAGTTGTGGGCTTATTTAGGCTTGG-TAACGAAGCTTAACG 835  
Db 791 GCGTGTAAACGATGCTGATTTGAGAGTTGTTCCCTTGAGAGTGGCTTCCGAGCTTAACG 850  
QY 836 CGTGAAGTTGACCGGCTGGGAGTACGCTCGCAAGATTAAACCTCAAGATTTGACGAG 895  
Db 851 CGTTAAATCGACCGCTGGGAGTACGCGCGCAAGTTAAACCTCAAGATTTGACGAG 910  
QY 896 GACCGGCAACAGCGGTGATTAATGATTAATTTGATGCAACGGCAAAACCTTAACCTA 955  
Db 911 GCGCCGCAACAGCGGTGAGCATGTGTTAAATGATGCAACGGCAAAACCTTAACCTA 970  
QY 956 CCGTTGACATGATGAGGAATTTCTAGAGTAGATTAGTG-CTTGGGGAACGCTAACACAG 1014  
Db 971 CTCTTGACATCCAGGAATTTAGCAGAGATGCTTAGTGCTTCGGGAACCTTGACACAG 1030  
QY 1015 GTGCTGATGCTGCTGACGCTCGTCTGAGAGATGTTGGGTTAAGTCCGCAACGAGC 1074  
Db 1031 GTGCTGATGCTGCTGCTGACGCTCGTCTGAGAGATGTTGGGTTAAGTCCGCAACGAGC 1090  
QY 1075 GCAACCTTGTGCTAATTTGCCATCA--TTTGGTTGGGCACTTTAATGAGACTGCGGTG 1132  
Db 1091 GCAACCTTATCTCTTGTGTCAGACGCTAATGTGGGAACTCAAGGAGACTGCGGTG 1150  
QY 1133 ACAAACGCGGAGAGGTGGGAGTGAAGTCAAGTCCCTCAAGGCGCTTATGGGAGGCTTC 1192  
Db 1151 ATAAACCGGAGAGGTGGGAGTGAAGTCAAGTCAAGTCCCTTATGAGAGGCTTAC 1210  
QY 1193 ACACGTAATACATGAGCGCTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAG 1252  
Db 1211 ACAAGTGTCTAATGCGCATATCAAAAGAGGCAACTCGGAGAGCCAGGGAATCAT 1270  
QY 1253 AAAAGCGGTGATGTCGAGTGGAGTCTGCAACTGCACTCGTGAAGTCCGAATCGCTA 1312  
Db 1271 AAAAGCGGTGATGTCGAGTGGAGTCTGCAACTGCACTCGTGAAGTCCGAATCGCTA 1330  
QY 1313 GTAATCGGAGTACACATGTGCGCGTGAATCGTCCGCGGCTTGTACACACGCGCGT 1372  
Db 1331 GTAATCGTGAAGTCAAAATGCTACGCTGAATACGTTCCGCGCTTGTACACACGCGCGT 1390  
QY 1373 CACAACATGAGAGTGGGTTTCAACGAAGAGGTAGTGAACCGTAAAGAGGCGCTTGC 1432  
Db 1391 CACAACATGAGAGTGGGTTTCAACGAAGAGGTAGTGAACCGTAAAGAGGCGCTTAC 1450  
QY 1433 CACGCTGAGATTGATGACTGGGGTG 1457  
Db 1451 CACTTGTGATTGATGACTGGGGTG 1475

RESULT 13  
US-10-831-286A-48671  
; Sequence 48671, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG

APPLICANT: CHATELLIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILDRESS, DARRELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; PRIOR FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 48671  
; LENGTH: 1530  
; TYPE: DNA  
; ORGANISM: Citrobacter Youngae  
US-10-831-286A-48671  
Query Match 69.2%; Score 1008.4; DB 11; Length 1530;  
Best Local Similarity 82.4%; Pred. No. 5.3e-305;  
Matches 1206; Conservative 0; Mismatches 251; Indels 7; Gaps 4;  
QY 1 ATTGAACGCTGGCGGCAATGCTTTACATGCAAGTCCGAACGCGACGAGATGCTTG-- 57  
Db 22 ATTGAACGCTGGCGGCAAGCTTTACATGCAAGTCCGAACGCGAGAGAGCTTGC 81  
QY 58 -CATCTGCTGCGAGTGGCGGAGCGGTGAGTAATGCAATCGGAACGTATCCAGAAAGAGGG 116  
Db 82 TCCTTGGGTGACGAGTGGCGGAGCGGTGAGTAATGCTTGGGAACTCCGATGAGAGGG 141  
QY 117 GGTAAACGATCGAAGAAAGTGTCTAATACCGCATTACTTAAGGAGGAAAGAGCGGATC 176  
Db 142 GATTAACGATCGAAGAAAGTGTCTAATACCGCATTACTTAAGGAGGAAAGAGCGGATC 201  
QY 177 GAAAGACCTTGGCGCTTTTGGAGCGCGCGATGCTGATTAAGCTAGTTGGTGGGTTAAAGC 236  
Db 202 TTGGGGCTCTTGGCATCGAGTGTGCGCAGATGGAGTTAGTGGTGGGTTAAAGC 261  
QY 237 CTACCAAGCGGAGCATAGTATGTTGTTGTTGAGAGAGACGACGACCACTGGGACTAGA 296  
Db 262 TCACCTAGGCGGACATCTCTAGCTGTGTAAGAGATGACACGACACTGGAACCTAGA 321  
QY 297 CACGGCCGACCTCTACGCGGAGGAGCAGAGTGGGGAATTTTGGCAATGGGCGCAACCT 356  
Db 322 CACGCTCAAGCTCTTACGCGGAGGAGCAGAGTGGGGAATTTTGGCAATGGGCGCAACCT 381  
QY 357 GATCAGCAATGCGCGCTGAGTGAAGAGCCTTGGGTTGAAGCTCTTTCAGTGCAG 416  
Db 382 GATCAGCAATGCGCGCTGATGTAAGAGGCTTGGGTTGAAGCTTTCAGTGCAG 441  
QY 417 AAGAAAGGTTACGCTAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 476  
Db 442 AAGAAAGGCTTGAAGTAAATTAATCTCAGCAATGACCTTCTGCAAAAGAGCAGG 501  
QY 477 TAACTAGTGGCAGCAGCGCGGTAAATGAGGAGTCAAGCGTTAATCGGAATTAATCTG 536  
Db 502 TAACTAGTGGCAGCAGCGCGGTAAATGAGGAGTCAAGCGTTAATCGGAATTAATCTG 561  
QY 537 GCGTAAAGGAGTGCAGAGCGGCTTGTAGTCAAGTGAATTCGCCGGCTTAACTGG 596  
Db 562 GCGTAAAGGAGTGCAGAGCGGCTTGTAGTCAAGTGAATTCGCCGGCTTAACTGG 621  
QY 597 GAATTGCGTGTAACTACAAAGGCTAGAGTGTGGCAGAGGAGTGGAAATTTCCATGTTGA 656  
Db 622 GAATTGCGTGTAACTACAAAGGCTAGAGTGTGGCAGAGGAGTGGAAATTTCCATGTTGA 681  
QY 657 GCAAGTAAATGCGTAGATATGAAAGAACATCGATGGCGAAGGACGCTCTGGGTTAA 716  
Db 682 GCGGTAAATGCGTAGATATGAAAGAACATCGATGGCGAAGGACGCTCTGGGTTAA 741  
QY 717 CACTGACGCTCATGCAAGAACGTTGGGAGCAAAACGATTTAGTATCCCTGTGTTGCA 776  
Db 742 GACTGACGCTCATGCTGGAAGAGCTGGGAGCAAAACGATTTAGTATCCCTGTGTTGCA 801







Db 1282 GATATGTCGATCCGAGATGAGTGTGCAACTGCACTCATGAACTCGGAATCGGTAGTA 1341  
|  
Qy 1316 ATCGCGATCAGCATGTGCGCGGTGAATAGTCCGCGGTCTGTACACACCGCCGTAC 1375  
|  
Db 1342 ATCGTGATCAAAATGCAACGGTGAATAGTTCGCGGCTGTGACACCGCCGTAC 1401  
|  
Qy 1376 ACCATGGAGTGGTTTCAACGAAAGCAGTAGTCTAACCGTAAGAGGGCGCTTGCA 1435  
|  
Db 1402 ACCATGGAGTGGTTGCAAAAGAGTAGTACTTAACCTTCGGAGGGCGCTTAACAC 1461  
|  
Qy 1436 GGTGAGATTCACTGAGGCTG 1457  
|  
Db 1462 TTTGTGATTCACTGAGGCTG 1483  
|  
RESULT 16  
US-10-831-286A-48675  
; Sequence 48675, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILDRESS, DARRELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; PRIOR FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 48675  
; LENGTH: 1495  
; TYPE: DNA  
; ORGANISM: Proteus penneri  
US-10-831-286A-48675  
Query Match 69.0%; Score 1005.4; DB 11; Length 1495;  
Best Local Similarity 82.7%; Pred. No. 4.6e-304;  
Matches 1211; Conservative 0; Mismatches 246; Indels 8; Gaps 5;  
Qy 1 ATTTGAAGCTGGCGGCAATGCTTTTACATCATGCAAGTCAAGGCAAGCAGC--GATGCTTGC 58  
|  
Db 1 ATTTGAAGCTGGCGGCGGCGCTTACACATGCAAGTCAAGGCAAGGAAAGCTTGC 60  
|  
Qy 59 AT--CTGGTGCGAGTGGCGGAGCGGTGAGTAATGATCGGAACGTATCGAAGAGGG 116  
|  
Db 61 TTTCTTCTCTACGACGCGCGGAGCGGTGAGTAATGTAATGCGGANTCTGCCGATAGAGGG 120  
|  
Qy 117 GGTAAACCATCGAAGATGTGCTAATAACCGCATACTCTTAAGAGAGAAACAGGGGATC 176  
|  
Db 121 GATTAACTAAGTGAAGACGGTGGCTAATACCGCATGACGCTTAACGCAACCAAGAGGGGCTC 180  
|  
Qy 177 GAAAGACCTTGGCTTTTGAAGCGCGCGGATGTCTGATTAAGTAATGTGGGGTGAAGGC 236  
|  
Db 181 TTCGACCTTGGCTTATCGATGAAGCCCATATGGGATTAAGCTAATGAGTGAAGGAGC 240  
|  
Qy 237 CTACCAAGGCGAGCATGATGATGCTGAGAGGAGCAGCAGCAGCAGCAGTGGAGTGA 296  
|  
Db 241 TCACCTAGGCGACATCTCTAGCTGTGTGAGAGGATGATGACGACACACTGGGAGTGA 300  
|  
Qy 297 CACGCGCCAGACTCTTACGAGAGGAGCAGAGTGGGAAATTTTGAACAATGGCGCAAGCT 356  
|  
Db 301 CACGCGCCAGACTCTTACGAGAGGAGCAGAGTGGGAAATATTGCAACAATGGCGCAAGCT 360  
|  
Qy 357 GATCAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTGTTAAAGCTCTTCACTAGTGA 416  
|  
Db 361 GATCAGCAATGCGCGGTGAGTGAAGAGGCTTGAAGGTTGTAAGTCTTTCAGCGGG 420  
|  
Qy 417 AAGAAAGGTTACGGTAATATGCTGACTCATGACGCTATCGACAGAAAGCAACCGG 476

Db 421 AGAAAGGATTAAGTTAATACCTTTATCAATTAACCTTACCCGACAGAAAGACACCGG 480  
|  
Qy 477 TAACTACGTGCGACAGCGCGGTAAATCTGATGGGTGCAACCGTTAATCGGAATTAAC 536  
|  
Db 481 TAACTCCGTGCGACAGCGCGGTAAATCGAGGGGTGCAACCGTTAATCGGAATTAAC 540  
|  
Qy 537 GCGTAAAGGTCGCGAGCGGCTTTGTAAGTCAGATGGAATCCCGGGCTTAACCTG 596  
|  
Db 541 GCGTAAAGGTCGCGAGCGGCTTAATGATGATGGAATCCCGGGCTTAACCTG 600  
|  
Qy 597 GAATTCGCTTGAACCTCAAGGCTAGAGTGGCAGAGGAGTGAATTCATGTGTA 656  
|  
Db 601 GAATTCGATCTGAACCTGTTGGCTAGAGTCTTGAAGAGGGGGGTGAATTCATGTGTA 660  
|  
Qy 657 GCAAGTAAATGCGTGAAGATATGGAAGACATCGATGGCGAAGGAGGCTCTCGGTAA 716  
|  
Db 661 GCGGTAAATGCGTGAAGATATGGAAGACATCGATGGCGAAGGAGGCTCTCGGTAA 720  
|  
Qy 717 CACTGACGCTCATGCAAGAAAGCGTGGGAGCAACAGAAATTAATACCTGTAGTCCA 776  
|  
Db 721 GACTGACGCTCATGCGAAGAGCGTGGGAGCAACAGAAATTAATACCTGTAGTCCA 780  
|  
Qy 777 GCGCCTAAACGATGTC-AACTAGTTGTGGGCTTATTAGGCTTGTGAACGAACTAACG 835  
|  
Db 781 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
|  
Qy 836 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 895  
|  
Db 841 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
|  
Qy 896 GACCGGCAACAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 955  
|  
Db 901 GCGCGGCAACAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 960  
|  
Qy 956 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014  
|  
Db 961 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
|  
Qy 1015 GTCGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1074  
|  
Db 1021 GTCGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
|  
Qy 1075 GCAACCTTGTCAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1132  
|  
Db 1081 GCAACCTTGTCAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1140  
|  
Qy 1133 ACAACCGAGAGAGGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1192  
|  
Db 1141 ATTAACCGAGAGAGGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
|  
Qy 1193 ACACTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1252  
|  
Db 1201 ACACTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
|  
Qy 1253 AAAAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1312  
|  
Db 1261 AAAAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
|  
Qy 1313 GTATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1372  
|  
Db 1321 GTATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
|  
Qy 1373 CACACCATGAGAGTGGTTTCAACAGAGCAGGATGATTAACCGTAAAGAGGCGCTTGC 1432  
|  
Db 1381 CACACCATGAGAGTGGTTTCAACAGAGCAGGATGATTAACCGTAAAGAGGCGCTTGC 1440  
|  
Qy 1433 CACGTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
|  
Db 1441 CACGTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1465

RESULT 17

```

US-10-831-286A-48669
; Sequence 48669, Application US/10831286A
; Publication NO. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELLIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; PRIOR FILING DATE: 2003-04-24
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 48669
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Cedecea neteri
US-10-831-286A-48669

```

Query Match	69.0%;	Score 1005.2;	DB 11;	Length 1494;
Best Local Similarity	82.2%;	Pred. No. 5.3e-304;		
Matches 1204;	Conservative	0;	Mismatches 253;	Indels 7;
				Gaps 4

Qy	ATTGAAGCGCTGGGGGACATGCTTTACACATGCAAGATGMAACGGCACACAGATGGTGG---	57
Db	ATTGAAGCGCTGGGGGACAGCCCTTAACACATGATGTCAGCGGTACACGGGAGACTTGGC	60
Qy	58 -CATCTGGTGCAGATGGCGGACGGGTGAGTATGATCGAATCTATCCGAAGAGAGGG	118
Db	TCCTCTGGGTGACGAGCGGGCGGACGGGTGAGTATGTCGGGGATCTGGCTGATGAGAGGGG	120
Qy	117 GGTAAACGCAATCGAAGAATGTGCTAATAACCGCATATACTTAAAGAGGAAACAGGGGATC	176
Db	121 GATTAACTAAGTGAAGAAACGGTATGCTAAATCCGCAATACGTCCGCAAGCCAAAGAGGGGAGCC	180
Qy	177 GAAGAAGACTTGGCGCTTTTGGAGCGGCGGAGATGCTGATTAGCTAGTATGGTGGGTAAAGGC	236
Db	181 TTGGGCGCTCTTGGCCATCAGATGTAACCCAGATGGAGTTTACTGATGATGGTGAAGTATGGC	240
Qy	237 CTACCAAGGCGACGATCAGTATGTTGGTCTGAGAGGACGACAGCCACACTGGGACTAGA	296
Db	241 TCAACTAGGAGAGATCCCTAGTGGTCTGAGAGGTATGACAGCCACACTGGAATCTAGA	300
Qy	297 CACGGCCCAGACTTCTTACGGAGAGCGACATGGGGAAATTTTGGACATGGCGCGAAGCTT	356
Db	301 CACGGTCCAACTCTTACGGGAGGCGACAGTGGGAAATTTTGGCAATAGGGCGCAAGACTT	360
Qy	357 GATCCAGCAATGCGCGGTGATGGAAGAAAGGCGCTTGGGTTGAAAGCTCTTCACTGAG	416
Db	361 GATGACGCCATGCGCGGTGATGGAAGAAAGGCGCTTGGGTTGAAAGTCTTTCAGCGAG	420
Qy	417 AAGAAAAGGTTATCGGTAAATTAATCTGATCTATGACGGTATCGACAGAAAGACACCGGC	476
Db	421 AAGAAAGCGTTAAGTTAATTAATCTTAAAGGATTAAGTTCGAGAAAGAAAGACCGGC	480
Qy	477 TAACTACGTGCCAGACAGCCGCGGTAAATCGTAGGGTGCAGAGCGTTAATCGGAATTACTGG	536
Db	481 TAACTCCGTGCCAGACAGCCGCGGTAAATGCGAGGGTGCAGAGCGTTAATCGGAATTACTGG	540
Qy	537 GCGTAAAGGTTGGCGAGCGGCGCTTGTATGATCGAATGTGAATCCCGGCGTTAACTGG	596
Db	541 GCGTAAAGGCGACGACAGCGCGTTTGTAAATGCGAATGTGAATCCCGGCGTTAACTGG	600
Qy	597 GAATTGCGTTTGAATCTACAGGCTATGAGTGTGCGAGAGGAGGTGAAATCTCAATGTGTA	656
Db	601 GAATGCAATTCGAAATCTGCGACGCTTGAATCTTGTAGAGGGGGGTGAATCTCAAGGTGA	660
Qy	657 GCACTGAATGCGTATGAGATATGGAAGAACATCGATGCGAAGCGCAAGCTCTCTGGGTTAA	716

Db	661	GGGGGGAATGCGTGAAGATCTGAGGAATACCGGTGGGAAAGGGGCCCCCTGGACAA	720
Qy	717	CAC TAGCCTCATGCAAGAAACCGTGGGGAGCAAAACGAGTTAGATACCTTGTAAGTCCA	776
Db	721	GACTACCGCTCGAGGTCGGAAGCCGTGGGGAGCAAAACGAGTTAGATACCTTGTAAGTCCA	780
Qy	777	GGCCCTAAACGATGTCAACTAGTTGTGTGGGCGCTTATTAAGGCTTGG-TAAAGAAAGCTAAC	835
Db	781	CGCCCTAAACGATGTCAACTAGTTGTGTGGGCGCTTATTAAGGCTTGGCTTCCGAGCTAACG	840
Qy	836	CGTGAAGTTGACCGGCTGGGGAGTACGGTCGCAAGATTAAACTCAAAAGAAATTGACGGG	895
Db	841	CGTTAAGTTCGACCGCTGGGGAGTACGGGCGCAAGATTAAACTCAAAAGAAATTGACGGG	900
Qy	896	GACCCGCGCAACAGCGGTGATTAATGTGAGTTAATTTGATGCAACGCGCAAAAACCTTAACCTA	955
Db	901	GGCCCGCGCAACAGCGGTGAGTATGTGTATTATTTGATGCAACGCGCAAAACCTTAACCTA	960
Qy	956	CCCTTGACATGTAGGGAATTTTCTAGAGTATGATTAGTG-CTTGGGGAACGGCTAACACAG	1011
Db	961	CTTTTGACATCCAGAGAACTTTCCAGAGTGAATTTGTGCTTCCGGAACCTTGAGACAG	1021
Qy	1015	GTGCTGACATGCGTGTGTCAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGC	1071
Db	1021	GTGCTGACATGCGTGTGTCAGCTCGTGTGTTGAATGTTGGTTAAGTCCCGCAACGAGC	1081
Qy	1075	GCAACCCCTTGTCATTAATTGCCATC-ATTGGTTGGGCACTTTAATGAGATCTGCCGCTGA	1133
Db	1081	GCAACCCCTTATCTTTGTGTGGCCAGGGGTTCCGCGCAACCTCAAGAGACTGCAAGTGA	1144
Qy	1134	CAAAACGGAGGAAGGTGGGGAGTGAAGTCAAGTCCCTCATGAGCCCTTATGGGTTAGGGCTTCA	1197
Db	1141	TAAATCGAGGAAGGTGGGGAGTGAAGTCAAGTCAATCATATGCGCTTATCGAGTGGGCTTACA	1201
Qy	1194	CACGTAATACAATGCGCGGTACAGAGGGTTGCAACCCGCGAGGGGAGCTAATCTCAGA	1257
Db	1201	CACGTCATACAATGCGCGCATACAAGAAAGGAGCCTCGCGAGACAAAGCGGACCTCATTA	1266
Qy	1254	AAGCGCGTCGTAAGTCCGGAATCGAGATCTGCAACTCGGTGAAGTTCGAAATGGCTGAG	1311
Db	1261	AAGTCGTCGTAAGTCCGGAATGGAAATCTGCAACTGACCTCATGAAGTTCGAAATGGCTGAG	1321
Qy	1314	TAAATCGCGAGTACAGCATGTCCGCGGTGAATACGTTCCCGGGTCTTTATACACACCGCCGTC	1372
Db	1321	TAAATCGTAGATCAGAAATGCTACGAGTAACGTTATCGTTCCCGGGCTTTATACACACCGCCGTC	1381
Qy	1374	ACACCATGGAGTGGGTTTCAACAGAACAGGTAATCTAACCGTAAGAGGGCGGCTTGCC	1433
Db	1381	ACACCATGGAGTGGGTTTCAAAAGAAATAGGTAGCTTAACCTTGGAGGGCGGCTTACC	1440
Qy	1434	ACGGTAGAATTCATGACTGGGGTG	1457
Db	1441	ACTTGTGATTCATGACTGGGGTG	1464

```

RESULT 18
US-10-831-286A-48665
; Sequence 48665, Application US/10831286A
; Publication NO. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174_001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; PRIOR FILING DATE: 2003-04-24
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: PatentIn version 3.2

```

```
; SEQ ID NO 48665
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Cedecia davisae
US-10-831-286A-48665
```

Query Match	69.0%;	Score 1005.2;	DB 11;	Length 1503;
Best Local Similarity	82.2%;	Pred. No. 5.3e-304;		
Matches 1204;	Conservative	0;	Mismatches 253;	Indels 7; Gaps 4

Oy	1	TTTGAACGCTGGCGGATGCTTTAAACAATGCAAGTCGAAACGGCAGACGGAAGCTTGCA-	59
Db	2	ATTGAACGCTGGCGGAGGCTTAACATATGCAATGCAAGCTTAACAGGATGACTTGC	61
Oy	60	---TCGTGTGCGAAGTGGCGGACGGGTGAGTAATGATCGAACGTTATCAGAAAGGGG	116
Db	62	TACTCGGTGACGAGCGGGCGGACGGGTGATGATGTCCTGGGATCTGCCTGATGAGGGG	121
Oy	117	GSTAAACGATCGAAAGATGCTTAATACCGCATTACTTAAAGAGGAAAGCAGGGGATC	176
Db	122	GATTACTACTGGAACCGTAGCTAATACCGCATACGTGCAAGACCAAGAGGGGGAAC	181
Oy	177	GAAAGACCTTGGGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTGGTGGGTTAAAGC	236
Db	182	TTTCGGGCTCTTGGCCATCGAATGAACCCAGATGGATTAGCTATAGTGGGTATATGC	241
Oy	237	CTACCAAGCGCAGATCACTAGTGTGTCTGAGAGACGACACGCCACTGTGGACTGAGA	296
Db	242	TCACCTAAGCGACGATCCTTAGCTGTGTGAGAGATGACACGCCACTGGAATGAGA	301
Oy	297	CACGGCCGACACTCTTACCGGAGGACGACAGTGGGAAATTTTGGACAAATGGGCGCAAGCT	356
Db	302	CACGGTCCGACACTCTTACCGGAGGACGACAGTGGGAAATTTGCAATATGGGCGCAAGCT	361
Oy	357	GATCAGACAAATCCGCGTAGTGAAGAAAGCCCTTGGGTTGTAAAGCTCTTTCAGTCGAG	416
Db	362	GATGACGCCATGCGCGGTGTGTAAGAAAGCCCTTGGGTTGTAAAGCACTTTCAGGAGG	421
Oy	417	AAGAAAAGGTTACGGTAAATTAATCTGACTCATGACGGTATCGACAGAAAGACACCGGC	476
Db	422	AGGAAGGCAATTAAGTTAATTAACCTCGGTGATTGACGTTACTCGAGAAAGACACCGGC	481
Oy	477	TAACTACGTGCCAGACGCCCGGTAAATAGTAGGGTGGCAAGCTTAATCGGAATTACTGG	536
Db	482	TAACTCCGTGCCAGACGCCCGGTAAATAGTAGGGTGGCAAGCTTAATCGGAATTACTGG	541
Oy	537	GGGTAAAGGTCGGCAGCGGGGCTTGTAGTGCAGATGCGAAATCCCGGGCTTAACCTGG	596
Db	542	GGGTAAAGGTCAGCAGCGGGGCTTGTAGTGCAGATGCGAAATCCCGGGCTTAACCTGG	601
Oy	597	GAAATGCGTTTGAATCTACAAGGCTAGAGTGTGGCAGAGGAGAGTGGAAATTCATGATGA	656
Db	602	GAACTGCATTCCAAATCTGCCAAGCTTAGTCTGTAGAGGGGGGTGAATTCAGGGTGA	661
Oy	657	GCAGTGAATGCGTAGAGATAGAAAGACATGATGGCGAAGCGACGCTCTCGGGTTAA	716
Db	662	GGGTGAATATGCTAGAGATCTGAGAGAAATACCGGTGGCCAAAGCGGCCCTCCGACCAA	721
Oy	717	CACCTGACGCTCAACGAAAGGCTGGGGAGCAAAACAGATTATGATTAACCTGGTAAATCCA	776
Db	722	GACTGACGCTCAGAGTGGAAAGGCTGGGGAGCAAAACAGATTATGATTAACCTGGTAAATCCA	781
Oy	777	CGCCCTTAAACGATGCACTAGTGTGTGTGGGCTTATTTAGGCTTTGG-TAAAGAACTAAG	835
Db	782	CGCCCTTAAACGATGTCATTTGAGAGTGTGTGTGGGCTTATTTAGGCTTTGG-TAAAGAACTAAG	841
Oy	836	CGTGAAGTTGACCGCTGGGGAGTAGTGGTGGCAAGATTAAATCTCAAGAAATTTGACGGG	895
Db	842	CGTTAAGTCACCGCTGGGGAGTAGTGGTGGCAAGATTAAATCTCAAGAAATTTGACGGG	901
Oy	896	GACCCGCAAGAGGGGTGAATTATGTGAATTAAATTCATGACGACGCAAAAACCTTACTTA	955
Db	902	GGCCCGCAAGAGGGGTGAAGTGTGAATTAAATTCATGACGACGCAAAAACCTTACTTA	961

QY	956	CCCTTGA	CATGTAG	CCGAATTTTCTAG	AGATAGATTA	TGTG-CTTCGGGA	AGCCTA	CAACAG	101.4
Db	962	CTCTTGA	CATTC	CAAGAA	CACTTTCCAGAGATG	ATGTCCTTCGGGA	CTCTAG	CAGC	102.1
QY	1015	GTGCTG	ATGAGT	CGTGTG	CTGACTCGTGTG	AGATTTGGGTAA	GTGTC	CCGCAACAGC	107.4
Db	1022	GTGCTTGA	ATGAGT	CGTGTG	CTGACTCGTGTG	AGATTTGGGTAA	GTGTC	CCGCAACAGC	108.1
QY	1075	GCAAC	CCCTTGCAT	TAAATTG	CCATC-ATTGGTTGG	CACTTAA	TAGACTG	CCCGGTGA	113.3
Db	1082	GCAAC	CCCTTAA	CTTTGTG	CAAGCCGGTGTG	CGCGGGA	ACTCA	AAAGAGAC	114.1
QY	1134	CAAA	CCGAGAGAGT	GGGAGTGA	CGTCAAGT	CCTCA	TGCGCCCTTAA	TGGTAA	119.3
Db	1142	TAA	CTGAGAGAA	AGTGGGGAGTGA	CGTCAAGTCA	TAGCGCCCTTAA	GAGTAA	GGGACTCA	120.1
QY	1194	CACGTA	TATACA	TGGCCCGTCA	CAAGGGTTGCCA	ACCGCGAGGGGGAG	CAATCT	CAGA	125.3
Db	1202	CACGTG	CTACA	TATGGCCCGATCA	CAAAAGAGAAC	CGACCTCG	CGAGAGCA	AGCCGAC	126.1
QY	1254	AAGG	CGCGTGA	TCCGGATCGA	AGTCTGCA	ACTCGTGA	AGTGGAT	CGGATCG	131.3
Db	1262	AAGT	CGCGTGA	TCCGGATCGA	AGTCTGCA	ACTCGTGA	AGTGGAT	CGGATCG	132.3
QY	1314	TAA	TGCGGAGT	CAGCATGT	CGCGGTGA	TACGTTCCCGGAGT	CTTGTAC	ACACCGCCGCTC	137.3
Db	1322	TAA	TGTAAGAT	CAAGATGT	CTACGGGTAA	TACGTTCCCGGAGT	CTTGTAC	ACACCGCCGCTC	138.1
QY	1374	ACAC	CGATGGAGT	GGGTTTCA	CCAGAGC	AGATAGTCTAA	ACCGTAA	AGAGGGCGCTT	143.3
Db	1382	ACAC	CGATGGAGT	GGGTTTCA	CCAGAGC	AGATAGTCTAA	ACCGTAA	AGAGGGCGCTT	144.1
QY	1434	ACG	TGAATTCAT	GACTGGGGTG	1.457				
Db	1442	ACT	TGTAATTCAT	GACTGGGGTG	1.465				

```

RESULT 19
US-10-831-286A-48666
; Sequence 48666, Application US/10831286A
; Publication No. US20060046246A1
GENERAL INFORMATION:
APPLICANT: ZENG, QIANDONG
APPLICANT: CHATELLIER, SONIA
APPLICANT: MOIR, DONALD T.
APPLICANT: LACROIX, BRUNA
APPLICANT: CHILDRESS, DARELL
TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
FILE REFERENCE: 0329266-174.001
CURRENT APPLICATION NUMBER: US/10/831,286A
CURRENT FILING DATE: 2004-04-26
PRIOR APPLICATION NUMBER: 60/464,955
PRIOR FILING DATE: 2003-04-24
NUMBER OF SEQ ID NOS: 48788
SOFTWARE: Patentin version 3.2
SEQ ID NO 48666
LENGTH: 1514
TYPE: DNA
ORGANISM: Cedecia davisiae
US-10-831-286A-48666

```

Query Match	69.0%	Score 1005.2;	DB 11;	Length 1514;
Best Local Similarity	82.2%;	Pred. No. 5.3e-304;		
Matches 1204;	Conservative	0;	Mismatches 253;	Indels 7;
				Gaps 4;

QY 1 ATTGAACGCTGGCGGCGAGCTTTAACAATGCAAGTCGAAACGGCAGACGAGATGCTTGC 59  
 QY 1 ATTGAACGCTGGCGGCGAGCTTTAACAATGCAAGTCGAAACGGCAGAGATGCTTGC 60  
 Db 1 ATTGAACGCTGGCGGCGAGCTTTAACAATGCAAGTCGAAACGGCAGAGATGCTTGC 60  
 QY 60 ---TCTGTGTGCGAAGTGGCGGACGGGTGAGTAAATGCATTCGAAACGTAATCAGAAAGGGG 116  
 Db 61 TACTCGCTGACGAGCGGCGGACGGGTGAGTAAATGCTGAGGAATCTCCTGATGAGGGG 120

117 GGTAAAGCATGGAAGATGCTGTAATACCGCATATACCTTAAGAGAAAGCAGGGGATC 176  
121 GATTAATCTAGTGAACCGTAGTAATTAACCGCATTAACGTCGAAGACCAAAGGGGGAC 180  
177 GAAGAACCCTGGCTTTTGGAGCGGCCGATGTCGATTAAGTAGTGGTGGGTAAAGGC 236  
181 TTCCGGGCTCTTTCCATCGGATGGAACCCAGATGGGATTAAGTAGTGGTGGGTAAAGGC 240  
237 CTACCAAGGCGAGATCAGTAGTGGTGTGAGAGAGACGACCACTGAGACTGAGAG 296  
241 TCACCTAAGGCGAGATCCTAGCTGATGAGAGATGACCACTGAGACTGAGAG 300  
297 CACGGCCGAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAAATGGGGCAAGCT 356  
301 CACGGTCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAAATGGGGCAAGCT 360  
357 GATCCAGCAATGCGCGGTAGTGAAGAAAGGCTTCGGGTGTAAAGCTCTTTCAGTCCAG 416  
361 GATGACGCAATGCGCGGTAGTGAAGAAAGGCTTCGGGTGTAAAGCTCTTTCAGTCCAG 420  
417 AAGAAAGGTTACGGTAATTAATCGTACTCATGACGGTATCGACAGAAAGACCCGAC 476  
421 AGGAAGGCAATTAAGGTTAATACTTAGTGAATGACGTACCGCAGAAAGACCCGAC 480  
477 TAACTACGTCGAGAGCGCGGTAAATCGTAGGGTGCAGAGGTTAATCGGAAATTAAGT 536  
481 TAACTCCGTCGAGAGCGCGGTAAATCGTAGGGTGCAGAGGTTAATCGGAAATTAAGT 540  
537 GCGTAAGAGGTCGAGAGCGCGGTAAATCGTAGGGTGCAGAGGTTAATCGGAAATTAAGT 596  
541 GCGTAAGAGGTCGAGAGCGCGGTAAATCGTAGGGTGCAGAGGTTAATCGGAAATTAAGT 600  
597 GAATTCGCTTGAATCTACAGGCTAGAGTGGCAGAGGAGGTGGAATTCATGTGTA 656  
601 GAATTCGCTTGAATCTACAGGCTAGAGTGGCAGAGGAGGTGGAATTCATGTGTA 660  
657 GCAGTAAGTGGTGAATCTAGAGTGAAGAAATCGTAGGGTGCAGAGGTTAATCGGAAATTA 716  
661 GCAGTAAGTGGTGAATCTAGAGTGAAGAAATCGTAGGGTGCAGAGGTTAATCGGAAATTA 720  
717 CACTGACGCTCTGACGAGAAAGCGTGGGAGCAAGAGTGAATTCAGTGGTGAATTC 776  
721 GACTGACGCTCTGACGAGAAAGCGTGGGAGCAAGAGTGAATTCAGTGGTGAATTC 780  
777 GCGCTTAAACGATGTCATAGTGTGGGCTTAATAGGCTTGAATTCAGTGGTGAATTC 835  
781 GCGCTTAAACGATGTCATAGTGTGGGCTTAATAGGCTTGAATTCAGTGGTGAATTC 840  
836 CGTGAAGTGAACCGCTGGGAGTACCGTGGCAAGTAAATCAAGGAAATTCAGTGGTGA 895  
841 CGTGAAGTGAACCGCTGGGAGTACCGTGGCAAGTAAATCAAGGAAATTCAGTGGTGA 900  
896 GACCCGCAAGAGCGGTGATTAATGATGATTAATTCAGTGGTGAATTCAGTGGTGA 955  
901 GACCCGCAAGAGCGGTGATTAATGATGATTAATTCAGTGGTGAATTCAGTGGTGA 960  
956 CCCTTGAATGATGATTAATTCAGTGGTGAATTCAGTGGTGAATTCAGTGGTGA 1014  
961 CCCTTGAATGATGATTAATTCAGTGGTGAATTCAGTGGTGAATTCAGTGGTGA 1020  
1015 GTGCGTGAATGATGATTAATTCAGTGGTGAATTCAGTGGTGAATTCAGTGGTGA 1074  
1021 GTGCGTGAATGATGATTAATTCAGTGGTGAATTCAGTGGTGAATTCAGTGGTGA 1080  
1075 GCAACCCCTTGTATTAATTCAGTGGTGAATTCAGTGGTGAATTCAGTGGTGA 1133  
1081 GCAACCCCTTGTATTAATTCAGTGGTGAATTCAGTGGTGAATTCAGTGGTGA 1140  
1134 CAAACCGGAGAGGTTGGAGTGAATTCAGTGGTGAATTCAGTGGTGAATTCAGTGGTGA 1193  
1141 TAAACTGAGAGAGGTTGGAGTGAATTCAGTGGTGAATTCAGTGGTGAATTCAGTGGTGA 1200

1194 CACGTAATACATGAGCGGTACAGAGGTTGCCAACCAGGAGGGGACTAATCTCAGA 1253  
1201 CACGTCATCAATGAGCGCATACAAAGAGAGCAGCTTCGAGAGCAAGCGGACTCAT 1260  
1254 AAGCGCGTGTAGTCCGATCGAGTGTGCACTCGACTCCGTGAAGTCCGATCGTAG 1313  
1261 AAGTCGTCGTAGTCCGATCGAGTGTGCACTCGACTCCGTGAAGTCCGATCGTAG 1320  
1314 TAAATCGGAGTACAGTGTCCGATCGAGTGTGCACTCGACTCCGTGAAGTCCGATCGTAG 1373  
1321 TAAATCGGAGTACAGTGTCCGATCGAGTGTGCACTCGACTCCGTGAAGTCCGATCGTAG 1380  
1374 ACACCATGGAGTGGGTTTCCAGAGAGAGTGTGTAAACCTGAAGAGGCGCTTGGC 1433  
1381 ACACCATGGAGTGGGTTTCCAGAGAGAGTGTGTAAACCTGAAGAGGCGCTTGGC 1440  
1434 ACGGTGAGATTCATGACTGGGGTG 1457  
1441 ACTTGTGATTCATGACTGGGGTG 1464

RESULT 20  
US-10-831-286A-48676  
; Sequence 48676, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHARBELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; TITLE OF INVENTION: GENUS, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-114.001  
; CURRENT FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; PRIOR FILING DATE: 2003-04-24  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 48676  
; LENGTH: 1494  
; TYPE: DNA  
; ORGANISM: Enterobacter hormaechei  
US-10-831-286A-48676

Query Match 69.0%; Score 1004.8; DB 11; Length 1494;  
Best Local Similarity 82.7%; Pred. No. 7;le-304;  
Matches 1210; Conservative 0; Mismatches 247; Indels 7; Gaps 5;

1 ATGGAAGCTGGCGGCAATGCTTTACATGCAAGTGCAGACGAGAT--GCTTGC 58  
1 ATTGAAGCTGGCGGCAATGCTTTACATGCAAGTGCAGACGAGATGCTTGC 60  
59 ATCT--GATGCGAGTGGCGGAGCGGTGAGTGAATGATCGGAAGTCCAGAAAGAGGG 116  
61 TGCTTCGCTGACGATGCGGAGCGGTGAGTGAATGCTCGGAAATCTCGATGAGAGGG 120  
117 GGTAAAGCATGGAAGATGCTGTAATACCGATTAATCTTAAGAGAGAAAGCAGGGATC 176  
121 GATTAATCTAGTGAACCGTAGTAATTAACCGCATTAACGTCGAAGACCAAAGGGGAGCC 180  
177 GAAGAACCCTGGCTTTTGGAGCGGCCGATGTCGATTAAGTAGTGGTGGGTAAAGGC 236  
181 TTCCGGGCTCTTTCCATCGGATGGAACCCAGATGGGATTAAGTAGTGGTGGGTAAAGGC 240  
237 CTACCAAGGCGAGATCAGTAGTGGTGTGAGAGAGACGACCACTGAGACTGAGAG 296  
241 TCACCTAAGGCGAGATCCTAGCTGATGAGAGATGACCACTGAGACTGAGAG 300  
297 CACGGCCGAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAAATGGGGCAAGCT 356  
301 CACGGTCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAAATGGGGCAAGCT 360

357 GATCCAGCATGCCGCGTGAAGTGAAGAGGCGCTTGCGGTTGTAAAGCTCTTTCAGTCCAG 416  
361 GATGACGCGATGCCGCGTGTATGAAGAGGCGCTTGCGGTTGTAAAGTACTTTTCAGCGGG 420  
417 AGAAGAGGTTACGGTAATTAATCGTCACTCAATGACGATGACAGAGAGACCGGCG 476  
421 AGAAGGCGATAGATTAAATTAATCTCAGCAATTTGACGTTAACTCCGCAAGAGAGACCGGCG 480  
477 TAATACGTGCGACGAGCGCGGTAAATAGTAGGGTGAAGCGGTTAATCGGAATTAATCG 536  
481 TAATCCGTGCGACGAGCGCGGTAAATAGTAGGGTGAAGCGGTTAATCGGAATTAATCG 540  
537 GCGTAAAGGCGTGGCGAGCGCGCTTTGTAAAGTCAAGATGTGAATCCCGGGGCTTAACCTGG 596  
541 GCGTAAAGGCGACGAGCGCGCTCTCAAGTCCGAGTGAATCCCGGGGCTCAACCTGG 600  
597 GAATTCGCTTTAAATTAACAGGCTGAGAGTGGCGAGAGGGAGGTGAATTCATCTGTGA 656  
601 GAATGCAATTCGAATCTGCGAGGCTAGAGCTTTGTAGAGGGGGTTGAAATTCAGGTGTA 660  
657 GCAGTGAATTCGTAGAGATATGAAGAAATATCGATGGCGAAGCGCGCTCTGGGTAA 716  
661 GCGGTGAATTCGTAGAGATATGAAGAAATATCGATGGCGAAGCGCGCTCTGGGTAA 720  
717 CACTACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGATACCTGTGATGCCA 776  
721 GACTACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGATACCTGTGATGCCA 780  
777 CGCGCTTAACGATGTCACTAGTTGTGGGCTTATTAAGCTTGG--TAAAGAGGTAAG 835  
781 CGCGCTTAACGATGTCACTAGTTGTGGGCTTATTAAGCTTGG--TAAAGAGGTAAG 840  
836 CGTGAAGTTGACCGCGTGGGAGTACGATGCGTGAAGTAAATCTCAAGAGAAATTTGACGG 895  
841 CGTGAAGTTGACCGCGTGGGAGTACGATGCGTGAAGTAAATCTCAAGAGAAATTTGACGG 900  
896 GACCCGCAAGAGCGGTGATATGTGGATTAATTCGATGCAAGCGCAAAAACCTTAACCTA 955  
901 GCGCGCAAGAGCGGTGATATGTGGATTAATTCGATGCAAGCGCAAAAACCTTAACCTA 960  
956 CCCTTGAATGATGCGCAATTTTCTAGAGATTAAGT--CTTCCGGAAGCGTAAACAG 1014  
961 CTCTTGAATGATGCGCAATTTTCTAGAGATTAAGT--CTTCCGGAAGCGTAAACAG 1020  
1015 GTGCTGATGAGTGTGCTGACGCTGCTGCTGCTGATGATTTGGTTAAGTCCCGCAACGAGC 1074  
1021 GTGCTGATGAGTGTGCTGACGCTGCTGCTGCTGATGATTTGGTTAAGTCCCGCAACGAGC 1080  
1075 GCAACCTTGTCAATTAATGCAATC--ATTTGGTTGGCACTTTAATGATGATGCTGCGGTGA 1133  
1081 GCAACCTTGTCAATTAATGCAATC--ATTTGGTTGGCACTTTAATGATGATGCTGCGGTGA 1140  
1134 CAACCGAGAGAGAGTGGGATGATGCTCAAGTCTCTAGGCGCTTATGAGGTAGGCTTCA 1193  
1141 TAAATGAGAGAGAGTGGGATGATGCTCAAGTCTCTAGGCGCTTATGAGGTAGGCTTCA 1200  
1194 CAGGTATATGAATGGCGGTGACAGAGGTTGCCAACCGCGAGAGGGAGCTTAATCTCAGA 1253  
1201 CAGGTATATGAATGGCGGTGACAGAGGTTGCCAACCGCGAGAGGGAGCTTAATCTCAGA 1260  
1254 AAGCGCTGTAGTCCGATCGAGTCTGCACTCGATCGGAAGTGGGAATCGGTAG 1313  
1261 AAGTGGCTGTAGTCCGATCGAGTCTGCACTCGATCGGAAGTGGGAATCGGTAG 1320  
1314 TAATCGGATCAGCATGTGCGGTGAATAGCTTCCGGGTCTTGTACACAGCCCGCTG 1373  
1321 TAATCGGATCAGCATGTGCGGTGAATAGCTTCCGGGTCTTGTACACAGCCCGCTG 1380  
1374 ACACCAAGGAGTGGGTTTCAACGAAGCAGGTAGTGAACCGTAACGAGAGGGCGCTTACC 1433  
1381 ACACCAAGGAGTGGGTTTCAACGAAGCAGGTAGTGAACCGTAACGAGAGGGCGCTTACC 1440  
1434 ACAGTGAATTCATGACTGGGTG 1457

Db 1441 ACTTGTGATTCATGACTGGGGTG 1464  
RESULT 21  
US-10-831-286A-48678  
; Sequence 48678, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHIDRESS, DARELL  
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; PRIOR FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: Patencin version 3.2  
; SEQ ID NO: 48678  
; LENGTH: 1535  
; TYPE: DNA  
; ORGANISM: Proteus penneri  
US-10-831-286A-48678  
Query Match 68.9%; Score 1003.8; DB 11; Length 1535;  
Best Local Similarity 82.6%; Pred. No. 1.5e-303;  
Matches 1210; Conservative 0; Mismatches 247; Indels 8; Gaps 5;  
1 ATTGAACGTCGCGCGATCTTTTACATCATGCAAGTCAAGCGCAGAC--GATGCTTGG 58  
21 ATTGAACGTCGCGCGCGATCTTTTACATCATGCAAGTCAAGCGCGTAAAGAAAGCTTGC 80  
59 AT--CTGTGGGAGAGGCGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGAT 116  
81 TTTCTTGTGACGAGCGCGGAGCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGAT 140  
117 GGTAAACGTCGCGCGATCTTTTACATCATGCAAGTCAAGCGCAGAC--GATGCTTGG 176  
141 GATTAATCTGAAACGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 200  
177 GAAAGCCTTGGCGCTTTTGAAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 236  
201 TTGGAACCTTGGCGCTTTTGAAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 260  
237 CTACCAAGCGCGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 296  
261 TCACCTAAGCGCGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 320  
297 CACGCGCCAGATCTCTAAGCGGAGGAGCAGAGTGGGAAATTTTGAACATGAGCGCAAGCCT 356  
321 CACGCGCCAGATCTCTAAGCGGAGGAGCAGAGTGGGAAATTTTGAACATGAGCGCAAGCCT 380  
357 GATCAGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 416  
381 GATGAGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 440  
417 AAGAAAGGTTACGGTAATTAATCGTCACTCAAGAGGTTGCAAGAGAGAACACCGGC 476  
441 AAGAAAGGTTACGGTAATTAATCGTCACTCAAGAGGTTGCAAGAGAGAACACCGGC 500  
477 TAATCGTGGCGACGAGCGCGGTAAATAGTAGGGTGAAGCGGTTAATCGGAATTAATCG 536  
501 TAATCGTGGCGACGAGCGCGGTAAATAGTAGGGTGAAGCGGTTAATCGGAATTAATCG 560  
537 GCGTAAAGGCGTGGCGAGCGCGCTTTGTAAAGTCAAGATGTGAATCCCGGGCTTAACCTGG 596  
561 GCGTAAAGGCGTGGCGAGCGCGGTCAATTAAGTCAAGATGTGAATCCCGGGCTTAACCTGG 620  
597 GAATTCGCTTTGAATTAACAGGCTGAGAGTGGCAGAGGAGGTGAATTCATCTGTGA 656



Db	621	GAATTGCATCTGAAACCTGGTTGACTGTAAGTCTTTGTAAGAGGGGGGTAGAAATTCATGTGTA	680
QY	657	GCAGTGAATAATGCGTAGAGATATGAGAAACATCGATGCGGAAGGAGCCTCTCGGGTTAA	716
Db	661	GCGGTGAATAATGGGTAGATGTGTGAGGAATACCGGTGGCGGAAGGGGGCCCCCTGGACAA	740
QY	717	CAC TAGACCTCATGCA CGAAGCGTGGGGAGCAACAGATTAGATACCTTGTA GTCCA	776
Db	741	GACTGACGCTCAGGTGCGAAACGTTGGGAGCAACAGATTAGATACCTTGTA GTCCA	800
QY	777	CGCCCTAAACGATGTC -AACTAGTTGTTGGGCTTATTAGGCTTGTAACGAAGCTAACG	835
Db	801	CGCTGTAAACGATGTGCTGATTTAGAGGTGTGTCTTGAACCGTGGCTTCTGAGACTAACG	860
QY	836	CGTGAAGTTGACCCGCTGGGGAGTAGTACGTGCGCAAGTTTAAACTCAAGAAATTGACGGG	895
Db	861	CGTTAAATGACCCGCTGGGGAGTAGCGCGCAAGTTTAAACTCAAGAAATTGACGGG	920
QY	896	GACCCGCACAGCCGCTGGATTATGTGTGATTTAATTCGATGCAACGCGA AAACTTACCTTA	955
Db	921	GCGCCGCACAGCCGCTGGAGTAGATGTGTGTTAATTCGATGCAACGCGA AAACTTACCTTA	980
QY	956	CCCTTGACATGTAAGGAATTTCTGTAGATAGATTAGTGC -CTTCCGGAAACGCTTAACAG	1011
Db	981	CTCTTGACATGCAAGCGGATCTTTAGAGTAAAGAGAGTGCCTTCCGGAAACGCTGAGACAG	1044
QY	1015	GTGCTGACATGAGTGTGTGTCAGCTCGTGTGAGAGATGTTGGTTAAGTCCCGCAACGAGC	1074
Db	1041	GTGCTGACATGAGTGTGTGTCAGCTCGTGTGTTAAGTGGGTTAAGTCCCGCAACGAGC	1100
QY	1075	GCAACCCCTTGTCATTTAATTCGCATCACTTTG -GTTGGGCACTTTAATGAGATTCGCCGGT	1133
Db	1101	GCAACCCCTTATCTTTGTTGTCAGCGCGCTGATGCGGGGAATCAAGAGGACTGCGCGTG	1166
QY	1133	ACAAACCGGAGGAAGGTGGGAGTAGACGTCAAGTCCCTCAATGGGCCCTTATGGATAGGGCTTC	1192
Db	1161	ATAAACCGGAGGAAGGTGGGAGTAGACGTCAAGTCAATGACCTCTTAACGATAGAGGCTTAC	1222
QY	1193	ACACGTAATACATAGGCGCGGTACAGAGGGTTGCCAAACCCGCGAGGGGAGACTAATCTCAG	1255
Db	1221	ACAGGTGCTACATAGGCGAGATACAAAGAGAACGCACTTGGGAGAGCAAGCGGAACCTCAT	1280
QY	1253	AAAGCGCGCTCGTAATGTCGAGTCCGAGTCTGCAATCTGACTCCGTGAATGTCCGAAATCGCTA	1311
Db	1281	AAAGTCTGTCTGTATGTCGAGATTGGAGTCTGCAACTCGACTCATAGAAATGTCCGAAATCGCTA	1340
QY	1313	GTAATTCGGGAGTACAGATGTGCGCGGTGAATAGCTTCCGGGCTCTTGTAACAACGCGCGGT	1372
Db	1341	GTAATTCGTAATCAGAAATGCTACGGGTGAATACGTTCCGGGCTCTTGTAACAACGCGCGGT	1400
QY	1373	CACACCATGAGAGTGGGTTTCAACGAAGCAGATGACTTAAACGTTAAGAGGGGCGCTTGC	1433
Db	1401	CACACCATGAGAGTGGGTTTCAAAAGAAAGTATGATGCTTAACTTCCGGAGGGCGCTTAC	1466
QY	1433	CACGGTGAATTCATGACTGGGGGTG 1457	
Db	1461	CACTTGTGATTCATGACTGGGGGTG 1485	

```

RESULT 22
US-10-831-286A-48670/C
; Sequence 48670, Application US/10831286A
; Publication No. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174..001
; CURRENT APPLICATION NUMBER: US/10/831,286A

```

```

; CURRENT FILING DATE: 2004-04-26
;
; PRIOR APPLICATION NUMBER: 60/464,955
; PRIOR FILING DATE: 2003-04-24
;
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: PatentIn version 3.2
;
; SEQ ID NO 48670
;
; LENGTH: 1501
;
; TYPE: DNA
; ORGANISM: Cedecaea neteri
;
US-10-831-286A-48670

```

Query Match	68.9%	Score 1003.6;	DB 11;	Length 1501;
Best Local Similarity	82.2%;	Pred. No. 1.7e-303;		
Matches 1203;	Conservative 0;	Mismatches 254;	Indels 7;	Gaps 4;

Oy	1	TTTGAAGCCTGGCCGCGCATGCTTTTACATGTGAAGTCGAACGGGACGACGAGATGCTGG---	57
Db	1497	ATTGAACCTTGCCGGCAGGCGCTTAACATACATGTGAAGTCGAGCGGTACACAGGAGAGCTTCC	1433
Oy	58	-CATCTGTGCGCAGTGTGGCGGACGGGTGAGTAATGCATCGGAACGTATCCAGAAAGAGGG	116
Db	1437	TCTCTGGGTGAGAGCGCGCGGAGCGGTGAGTAATGTCTGGGAGTCTGCTGATGGAGGG	1376
Oy	117	GGTAAAGCATGGAAAGATGTGCTTAATACCGCATATCTTAAAGAGGAACAGAGGGATC	176
Db	1377	GATAACTCTGGAACCGGTACTTAATACCGCATTAACGTGCGMAACCAAAAGAGGGAGCC	1318
Oy	177	GAAGAACCCTTGCGCTTTTGGAGCGGCGCATGTCTGAATTAGCTAGTGGTGGGTAAAGGC	236
Db	1317	TTTCGGGCTCTTGGCCATCAGATGAACCCAGATGGGATTTAGCTAGTGGTGAAGTAATGC	1258
Oy	237	CTACCAAGCGACGATCAGTAAGTTGCTCTGAGAGAGACGACCGACACTGGGACTGAGA	296
Db	1257	TCACCTAAGCGACGATCCCTAGTGTGTCTGAGAGGATGACACGCACTCTGAACCTGAGA	1198
Oy	297	CACGGCCCAACTCTTACGGGAGCGACGATGGGGAAATTTTGACATATGGCGCAAGCT	356
Db	1197	CACGGTCCAGACTCTTACGGGAGCGACGATGGGGAAATTTGACATATGGCGCAAGCT	1138
Oy	357	GATCAGCAATGCGCGCTGAGTGAAGAAAGGCTTCGGGTGTAAAGCTCTTTCAGTCAG	416
Db	1137	GATGACGCAATGCGCGCTGATGAAGAAAGGCTTCGGGTGTAAAGTCTTTCAGTCAG	1078
Oy	417	AAGAAAGGTTACGGTAAATATATCGTGACTCATGACGGTATCGACAGAAAGACCCGGC	476
Db	1077	AGGAAGGATTAAGGTTAATAACCTTACGATGACGTTACTCGAGAAAGAACCCGGC	1018
Oy	477	TAACTACGTGCACGACGCCGGGTAAATCGTAGGGTGCAGGCTTAATCGGAATTAATCTGG	536
Db	1017	TAACTCCGTGCAGACGCCGGGTAAATCGGAGGGTGCAGGCTTAATCGGAATTAATCTGG	958
Oy	537	GCGTAAAGGGTGCAGGCGGCTTTGTAAATCGATGTGAATATCCCGGGCTTAAACCTGG	596
Db	957	GCGTAAACGACCGCAGGCGGTTGTTAAGTCGATGTGAATATCCCGGGCTCAACTCTGG	898
Oy	597	GAATTGCTTTGAAACTCAAGGCTAGAGTGTGCAGAGGGAGGTGGAATTCATGTGTGA	656
Db	897	GAACTGCAATCGAAACTGGAAGCTGAGTCTTGTGAAGGGGGGTGAATTCAGAGTGTGA	838
Oy	657	GCAGTGAATGCGTGAAGATATGGAAGAACTCGATGGCGAAAGGCAAGCTCTCTGGGTTAA	716
Db	837	GCGGTGAATGCGTGAAGATCTGGAAGGAATACCGGTGCGAAAGCGCGCCCTTGACAA	778
Oy	717	CACGTAGCCTATGACGAAAGCGTGGGGAACAAACAGATTAGTATCCCTGATAGTCCA	776
Db	777	GACTGACGCTCAGGTGCGAAAGCGTGGGGAACAAACAGATTAGTATCCCTGATAGTCCA	718
Oy	777	CGCCCTAAACGATGTCAACTAGTTGTTGGGCTTATTAGCTTGG-TAACGAACCTAACG	835
Db	717	CGCGTAAACGATGTGCACTTGGAAGGTTGTGCCCTTGAGGCGTGTGGCTTCGGAAGCTAACG	658
Oy	836	CGTGAAGTTGACCGGCTGGGGAGTACGGTCCGAAGATTAAATCTCAAGGAATTGACGGG	895



Db	657	CGTTAAGTCGACCCGCTGGGGAGTACGGCCGCAAGGTTTAAATCTCAATGAAATTGACGGG	598
Qy	896	GACCCGCGACAAGCGGATGATTAATGATTAATTCATGCAACGCGAAAACCTTACCTA	955
Db	597	GGCCCGGACAAAGCGGATGAGCATGTGGTTTAATTCGATGCAACGCGAAACCTTACCTA	538
Qy	956	CCCTTTGACATGTAGCGAATTTTCTTGAGATAGATTAGTG-CTTGGGGAGCGCTAACAG	1014
Db	537	CTCTTGACATCGAGGAACCTTCCAGATGATGATGGCTTGGGGAACCTTGAGACAG	478
Qy	1015	GTGCTGACATGGCTGCGTCAAGTCGATGAGATGTTGGGTTAAAGTCCCGCAACGAGC	1074
Db	477	GTGCTGACATGGCTGCGTCAAGTCGATGAGATGTTGGGTTAAAGTCCCGCAACGAGC	418
Qy	1075	GCAACCCCTTGTCATTAAATTTGCCATC-ATTGSGTTGGGCACTTAAATGAGATGCCGGTGA	1133
Db	417	GCAACCCCTTATCTTTGTTGGCAGGCGGTTCCGGCCGGGAACTCAAGGAGAGATGCGCAGTGA	358
Qy	1134	CAACCCGAGAAAGGTGGGGAGTGAAGTCAAGTCCATAGGCCCTTAATGAGATGGGCTCA	1193
Db	357	TAAATCGAGAAAGGTGGGGAGTGAAGTCAAGTCAATCAATGAGGCCCTTAACAGATGAGGCTCA	298
Qy	1194	CACGTAAATACATGCGCGGCTACAGAGGGTTGCCAACCCGCGAGGGGGAGCTAATCTCAGA	1253
Db	297	CACGTGCTAACATGGCGCATACAAAGAGAGGACCTCGCGAGACAAAGCGGACTTCATA	238
Qy	1254	AAAGCGCTGTAAGTCCGATCGGAATCTGCAACTGCACTCCGTGAAGTCGAAATGGCTAG	1313
Db	237	AAAGTCGCTGTAAGTCCGATCGGAATCTGCAACTGCACTCCATGGAAGTCGGAATGCTAG	178
Qy	1314	TAAATCGCGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTTTACACACCGCCGTC	1372
Db	177	TAAATGTAATCAAGAAATGCTACAGTGAAATACGTTCCCGGGCTTTTACACACCGCCGTC	118
Qy	1374	ACACCATGGAGTGGGTTTACACAGAGCAGGTAAGCTTAACCGTAAAGAGGCGCTTGCC	1433
Db	117	ACACCATGGAGTGGGTTTCAAAAAGAGTAACTTAACCTTCGGGAGGCGCTTGACC	58
Qy	1434	ACGGTGAATTCATGATCGGGGTG	1457
Db	57	ACTTTGTGATTCATACATCGGGGTG	34

```

RESULT 23
US-10-831-286A-48693
/ Sequence 48693, Application US/10831286A
/ Publication No. US20060946246A1
/ GENERAL INFORMATION:
/ APPLICANT: ZENG, QIANDONG
/ APPLICANT: CHATELIER, SONIA
/ APPLICANT: MOIR, DONALD T.
/ APPLICANT: LACROIX, BRUNA
/ APPLICANT: CHILDRESS, DARRELL
/ TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
/ FILE REFERENCE: 032796-174.001
/ CURRENT APPLICATION NUMBER: US/10/831.286A
/ PRIOR FILING DATE: 2004-04-26
/ PRIOR APPLICATION NUMBER: 60/464,955
/ PRIOR FILING DATE: 2003-04-24
/ NUMBER OF SEQ. ID NOS: 48788
/ SOFTWARE: PatentIn version 3.2
/ SEQ. ID NO 48693
/ LENGTH: 1527
/ TYPE: DNA
/ ORGANISM: Morganella morganii sp. morganii
US-10-831-286A-48693

```

Query Match	68.8%	Score 1002.8	DB 11	Length 1527
Best Local Similarity	82.5%	Pred. No. 3e-303		
Matches 1210	Conservative	0	Mismatches 247	Indels 9
			Gaps	5

QY 1 ATTGAACGCTGGCGGCATGCTTTA-CACATGCAAGTCGACCGGACGACGGA-----TGCT 55

Db	14	ATTGAACGCTGGCGGCGGACGGCTTAAGACATGCAAGTCGGGCGGTAAACAGGAGAAAGCTTG	73
Qy	56	TGCATCTGTGGTCGAGATGCGGSAAGGGTGAATATGCAATCGATATCCAGAAAGGG	115
Db	74	CTTCTCTGCTACGACGCGCGGACGGGTGAATATGATGGGGATCTGCTGATGCGGG	133
Qy	116	GGGTAAACGCATCGAAAGTGTCTAATATCCGCATATCTCTAAGGAGAAAGCAGGGGAT	175
Db	134	GGATTAATCTAGAAACGGTAGCTAATAACGCATAAATGTCCTTCGGACCAAGCCGGGGAC	193
Qy	176	CGAAAGACTTTCGCGCTTTTGGAGCGGCGGATGTCGATTTAGCTAGTGGTGGTTAAAG	235
Db	194	CTCGGGGCTTGCGCCCATCAATGAAACCCATATGGGATTAGCTAGTGGTAGTTAAACG	253
Qy	236	CCTACCAAGGCGAGATCATGAGTTGGTCTTGAGAGGACGACCAAGCTCATCGGACTGAG	295
Db	254	CTTACTTAGGCGAGATCCCTAGCTGGTCTTGAGAGATGATCAGCCACACTGGGACTGAG	313
Qy	296	AACAGGCCACAGCTCTTACGGGAGGCGACAGTGGGAAATTTTGGACAATGGGCGCAAGCC	355
Db	314	AACAGGCCACAGCTCTTACGGGAGGCGACAGTGGGAAATTTTGGACAATGGGCGCAAGCC	373
Qy	336	TGATCTACGACATGCGCGGTGATGAAGAGGCGCTTCGGGTGTAAAGCTCTTCAAGTCGA	415
Db	374	TGATGACACCATGCGCGGTGATGAAGAGGCGCTTCGGGTGTAAAGTACTTTCAATGTCG	433
Qy	416	GAAAGAAAGTTACCGTTAATATATCTGATCTCATGACGGTATCGACAGAAAGACACCGG	475
Db	434	GAGAAAGGTGGCAAGGTTAATTAACCTTGCAATTTGACGTTACGACAGAAAGACACCGG	493
Qy	476	CTTAATCACTGTCAGACAGCCGCGTAAATACGTAGGGTGCAGACGTTATCGGAATTACTG	535
Db	494	CTTAATCCGTGCAAGCACCGCGGTAAATACGAGAGGTGCAGACGTTATCGGAATTACTG	553
Qy	536	GCGCTAAAGGGTGGCGCAGCGGCTTTGTAATCAATGATGTAATCCCGGGCTTAAACCTG	595
Db	554	GCGCTAAAGCGCACGACGCGGTGATTGAGTCAGATGTGAATCCCGGGCTTAAACCTG	613
Qy	596	GGAAATGGGTTGAAACTCAAGGCTAGTGTGCGACAGGAGGGTGAATTCATGTCGT	655
Db	614	GGAAATGGGTTGAAACTCAAGGCTAGTGTGCGACAGGAGGGTGAATTCATGTCGT	673
Qy	656	ACGAGTGAATATGCGTAGAGATATGGAAGAACATCGATGGCGAGCGACTTCCTGGGTTA	715
Db	674	ACGAGTGAATATGCGTAGAGATATGGAAGAACATCGATGGCGAGCGACTTCCTGGGTTA	733
Qy	716	ACAATGACGCTCATGACAGAAAGCTGGGGAGCAAAACAGATTTAATACCCTGATGATCC	775
Db	734	ACAATGACGCTCATGATGCGAAAGCTGGGGAGCAAAACAGATTTAATACCCTGATGATCC	793
Qy	776	ACGCCCTAAAGATGTCATCTAGTTGTTGGGCTTAATTAGCTTGG--TAAAGAACTAAC	834
Db	794	ACGCCCTAAAGATGTCATCTAGTTGTTGGGCTTAATTAGCTTGG--TAAAGAACTAAC	853
Qy	835	GCGTGAAGTTGACCGCTGGGGAGTACGGTGCAGAAATTTAAACTCAAGAAATTTGACGG	894
Db	854	GCGTGAAGTTGACCGCTGGGGAGTACGGTGCAGAAATTTAAACTCAAGAAATTTGACGG	913
Qy	895	GGAACCCGCAAGCGGTGATTAATGTGATTTAATTCATGCAACCGCAAAAACCTTACTCT	954
Db	914	GGAACCCGCAAGCGGTGATTAATGTGATTTAATTCATGCAACCGCAAAAACCTTACTCT	973
Qy	955	ACCCCTGAATGATAGGAAATTTCTAGAGATGATTTAGTG--CTTGGGAAACGCTAACACA	1011
Db	974	ACCCCTGAATGATAGGAAATTTCTAGAGATGATTTAGTG--CTTGGGAAACGCTAACACA	1031
Qy	1014	GGTGTGCATGCGCTGTCAGCTCGTGTGAGATGTTGGTTAAGTCCCGCAACGAG	1077
Db	1034	GGTGTGCATGCGCTGTCAGCTCGTGTGAGATGTTGGTTAAGTCCCGCAACGAG	1093
Qy	1074	CGCAACCTTGTGATTAATTTGCATCATTTG--GTTGGGCACTTTAATGAGCTGCGGCT	1133
Db	1094	CGCAACCTTGTGATTAATTTGTGCACGCGCTATGATGCGGGAACCTCAAGAGACTGCGGCT	1155

QY	1132	GACAAA	CCGAGGAAGGTGGGGATGACGTCAAGTCTTCATGSCCTTTATGGGTAGGGCTT	1191
Db	1154	GATAAA <td>CCGAGGAAGGTGGGGATGACGTCAAGTCTTCATGSCCTTTATGGGTAGGGCTT</td> <td>1213</td>	CCGAGGAAGGTGGGGATGACGTCAAGTCTTCATGSCCTTTATGGGTAGGGCTT	1213
QY	1192	CACACGTAAATACAAATGGCGCGCTACAGAGGGTTGCGCAACCCCGAGGGGGAGCTAATCTCA		1251
Db	1214	CACACGTGCTCAATAGCGCTATATCAAAAGGAGGAGACCCCGAGGGGAGCTAATCTCA		1273
QY	1252	GAAGGCGCTCGTAGTTCGCGATCGAGCTTGCAACTCGACTCGGTGAAGTCGAAATCGCT		1311
Db	1274	TAAAGTACGTCGTAGTCCGGATTGGAGTCTGCAACTCGACTCGAATGAAATCGAAATCGCT		1333
QY	1312	AGTAATCGCGGATCAGCATGTCCGGGTGAATACCTTCCCGGCTTTGTATCAACCGGCCG		1371
Db	1334	AGTAATCGTAGATACAGAAATCTACGGGTGAATACCTTCCCGGCTTTGTATCAACCGGCCG		1393
QY	1372	TCACACCATGGAGGTGGGTTTCACAGAGGAGGTAGTCTAACCGTAAGAGAGGGCGCTTG		1431
Db	1394	TCACACCATGGAGGTGGGTTTCACAGAGGAGGTAGTCTAACCGTAAGAGAGGGCGCTTG		1453
QY	1432	CCACGGTGAAGTTCATGACTGGGGTG	1457	
Db	1454	CCACTTGTGATTCATGACTGGGGTG	1479	
RESULT 24				
US-10-831-286A-48679				
Sequence 48679, Application US/10831286A				
Publication No. US20060046246A1				
GENERAL INFORMATION:				
APPLICANT: ZENG, QIANDONG				
APPLICANT: CHATELIER, SONIA				
APPLICANT: MOIR, DONALD T.				
APPLICANT: LACROIX, BRUNA				
APPLICANT: CHILDRESS, DARRELL				
TITLE OF INVENTION: GENES, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES				
FILE REFERENCE: 032796-174.001				
CURRENT APPLICATION NUMBER: US/10/831,286A				
CURRENT FILING DATE: 2004-04-26				
PRIOR APPLICATION NUMBER: 60/464,955				
PRIOR FILING DATE: 2003-04-24				
NUMBER OF SEQ ID NOS: 48788				
SOFTWARE: Patentin version 3.2				
SEQ ID NO 48679				
LENGTH: 1486				
TYPE: DNA				
ORGANISM: Enterobacter hormaechei				
US-10-831-286A-48679				
Query Match 68.6%; Score 1000; DB 11; Length 1486;				
Beet Local Similarity 82.3%; Pred. No. 2.3e-302;				
Matches 1197; Conservative 0; Mismatches 250; Indels 7; Gaps 4;				
QY	11	GGCGGATCTTTACACATCAAGTCGAAACGGCAGACGGA---TGCTTGCATCTGGTG	66	
Db	1	GGCGGAGGCGCTTACACATCAAGTCGAAACGGTACAGAAACGACTTGCCTTTGGTG	60	
QY	67	GGGATGGCGGACGGGTGACTAATGCATCGAACGTATCCAGAAAGGGGGGTAAACGAT	126	
Db	61	AGGATGGCGGACGGGTGAGTAAATGTCGGGAAACTCTGATGAGGGGGATTAATTAAT	120	
QY	127	CGAAAGATGTGCTAATACCGCATATTAATCAAGGAGGAAAGCAGGGGATCGAAAGACTT	186	
Db	121	GGAAACGTAGCTAATACCGCATATTAATCGCAGACCAAGAGGGGACCTTCCGGCTTC	180	
QY	187	GGCTTTTGGAGCGCGCATGTCTGATTAAGTCTAGTTGTTGGGTAAAGGCTTACCAAGC	246	
Db	181	TTGCGATCGGAGTGTGCCAGATGAGGATTAAGTCTAGTGGGTAAAGGCTTACCAAGC	240	
QY	247	GACGATCAGTAGTTGTTCTGAGAGGACGACCAAGCACACTGGGACTGAGACAGGCTTAC	306	
Db	241	GACGATCCTTAGCTGTTCTGAGAGGATGACGACGACCACTGGAATGAGACACGCTTAC	300	

QY	307	ACTCTTAAGGAGGACAGATGGGGAAATTTTGGACAATGGGCGCAAGCCTTATCCAGCA	366
Db	301	ACTCTTAAGGAGGACAGATGGGGAAATTTTGGACAATGGGCGCAAGCCTTATCCAGCA	360
QY	367	TGCGCGGATGTAAGAAGGCGCTTCGGGTGTGTAAGCTTTTCAGTCCAGAAAGAAAGT	426
Db	361	TGCCCGGTGTGAAGAAAGCGCTTCGGGTGTGTAAGCTTTTCAGCGGGAGAAAGGGA	420
QY	427	TACGGTAATATCGTGACTCATGACGATATCGACAGAAAGACACCGGCTTAACCTAGTG	486
Db	421	TAAAGTTAATAACCTTGTTCGATTGACGTTACCCCGACAGAAAGACACCGGCTTAACCTCGTG	480
QY	487	CCAGCAGCGCGGGTAAATTCGTAAGGTTCGAACCGTTAATCCGAATTTACCTGGGCGTTAAAGG	546
Db	481	CCAGCAGCGCGGGTAAATTCGAAGGGTGAACCGTTAATCCGAATTTACCTGGGCGTTAAACG	540
QY	547	TGCGCAGGCGGCTTTGTATAGTCAGATGTGAATCCCGGGGCTTTAACTGGGAATTCGCTT	606
Db	541	CACCGACGGGCTCTGTCAAGTCGATGTGAATCCCGGGGCTTCAACCTGGGAATTCGATTT	600
QY	607	TGAAATCTAACAGGCTAGAGTGTGGCAGAGGGAGGTGCAATTCATGTGTACAGTGAAT	666
Db	601	CGAAATCGGCAAGGCTAGAGCTTTGTAGAGGGGGGTAGATTCACAGGTGTACCGGTGAAT	660
QY	667	GCGTAAGATATGGAAGAACATGCATGGCGAAAGGACGCTCGCGGTTAACACTGACGCT	726
Db	661	GCGTAAGATCTGAGAGAAATCCGGTGGCGAAAGGCGGCGCTCTGACAAAGACTGACGCT	720
QY	727	CATCACAAGAAAGCGTGGGAGCAAAACAGATTTAGATACCTCTGTATGTCACAGCCCTTAAC	786
Db	721	CAGGTGGAAACGCTGGGAGCAAAACAGATTTAGATACCTCTGTATGTCACAGCCGTTAAC	780
QY	787	GATGTCAACTATGTTGTTGGGCTTATTAAGCTTGG-TACGAAAGCTAACCGGTGAAGTTG	845
Db	781	GATGTGCACTTGGAGGTGTGTCCTTGGAGGGGTGGCTTCCGAACTAACCGGTTAAGTCG	840
QY	846	ACCGGCTGGGAGATACGGTGGCAAGATTAAACTCAAAAGAAATTTGACGGGAGACCGGACAC	905
Db	841	ACCGGCTGGGAGATACGGCCCGCAAGGTTAAACTCAAAAGAAATTTGACGGGAGACCGGACAC	900
QY	906	AGCGGTGATTTATGTGATTTAATTCGATGCAAGCGCAAAAACCTTACTCTTTCAT	965
Db	901	AGCGGTGATGATGTGTTAATTCGATGCAACCGCAAAAACCTTACTCTTTCAT	960
QY	966	GTAACGAATTTCTAAGATATGATTTAAGG-CTTGGGAAAGCTTAACACAGGTGTCGATG	1024
Db	961	CCAGAAGACTTACAGAGATGCTTTGGTGGCTTGGGAAACTTGAAGACAGGTGTCGATG	1020
QY	1025	GCTGTGCGACGCTCGTGTCTGAGATATGTTGGGTTAAGTCCCGCAACGACGCAACCTTTG	1084
Db	1021	GCTGTGCGACGCTCGTGTGTGTGAATATGTGGGTTAAGTCCCGCAACGACGCAACCTTTA	1080
QY	1085	TCATTTAATGTCATC-ATTGGTGGGCACTTTAATGACATCGCGGCTGACAAACCGGAG	1143
Db	1081	TCCTTTGTGTCACACGGTTAGCGCGGGAACCTCAAGGAGATGCGCAGATTAATACCTGGAG	1140
QY	1144	GAAAGTGGGATGTAACGTCAACTCTCATAGGCCCTTAAGGATAGGGCTTCAACGTAAATAC	1203
Db	1141	GAAAGTGGGATGTAACGTCAACTCATAGGCCCTTACAGATAGGACTTCAACACGCTGCTAC	1200
QY	1204	AATGGCGGTTCACAAGGGTTGCCAACCCGCGAGGGGGAGCTAATCTCGAAGAGCGGCTGC	1263
Db	1201	AATGGCGGTTCACAAGGAAGACGACTCGGAGAGCAAGCGGACCTCAATAAAGTCGTGC	1260
QY	1264	TAGTTCGAGATCGAGACTCTGCAACTCGACTCCGTAAGTCGGAATCGTAATTCGCGGA	1322
Db	1261	TAGTTCGAGATTCGAGACTCTGCAACTCGACTCAATCGGAATCGCTAATTCGTGGA	1320
QY	1324	TCACAGATGCGCGGTGAATACGTTCCCGGGCTTGTACACACCGCCGCTACACATATGG	1383
Db	1321	TCACAAATGCCAGGTGAATATCGTTCCCGGGCTTGTACACACCGCCGCTACACATATGG	1380

Qy	1384	AGTGGGCTTCACCGAAGCAGACTAGTGTCTAACCCGTAAAGAGGGCGCTGCCAGCGTGAT	1443
Db	1381	AGTGGGCTTCGAAAAGAAGTAGTAGTTAACCTTCGGAGGGCGCTTAACAATTGTGAT	1440
Qy	1444	TCATGACTGGGGTG	1457
Db	1441	TCATGACTGGGGTG	1454

RESULT 25  
US-09-941-095-158  
; Sequence 158, Application US/09941095  
; Publication No. US20060068383A1  
; GENERAL INFORMATION:  
; APPLICANT: BROW, MARY ANN D.  
; LYAMICHEV, VICTOR I.  
; OLIVE, DAVID M.  
TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF PATHOGENS  
NUMBER OF SEQUENCES: 165  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MEDLEN & CARROLL  
STREET: 220 MONTGOMERY STREET, SUITE 2200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/941,095  
FILING DATE: 28-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/941,193  
FILING DATE: 28-Aug-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, PETER G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: FONS-01756  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 158:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1542 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 158:

Query Match            68.3%; Score 995.2; DB 8; Length 1542;  
Best Local Similarity 82.2%; Pred. No. 7,3e-301;  
Matches 1204; Conservative 0; Mismatches 253; Indels 7; Gaps 5

Qy	1	ATTGAACGGTGGCGGCATCTTTACATCATGCAAGTCGAACGGCAAGCACG-GATGCTTGC	58
Db	28	ATTGAACGGCTGGGGCGACGCTTAACACATGCAAGTCGAACGGTTAACAGAGAACCTTGC	87
Qy	59	ATC--TGGTGGCAGTGGCGGACCGGTGAGTAATGATCGGAACGTATCCAAGAGGGGG	116
Db	88	TTCCTTTCGACGACGATGGCGGACGGGTGAGTAATGCTGGAAACTGCTGATGGAGGGG	147
Qy	117	GGTAAACGATCGAAATGCTATAATACCGCATATCTTAAAGAGAAAGCAGGGGATC	176
Db	148	GATTAATCTACTGAAACGTAAGCTAATAACCGCATTAAGTGGCAAGACCAAAGAGGGGACC	207
Qy	177	GAAGAAGCTTGGCGCTTTTGGAGCGGCGCATGTCTGATTAAGCTAGTTGGGGGTAAAGGC	236

D	b	208	TTGGGGCTCTGGCATCGGATGCGCCAGATGGGATTTAGCTAGTAGTGCGGTAACCGG	267
O	y	237	CTACCAAGGCGCATCAGTAGTGGTCTGAGAGACGACACGACACACTGGGGCTGAGA	296
D	b	268	TCACCTTAGGCGACATCCCTAGCGGTCTGAGAGATGACACGACACACTGGAATGAGA	327
O	y	297	CACGGCCCAAGCTCTTCGCGGAGGACACAGTGGGGAATTTTGGCAATGGGCGCAAGCT	356
D	b	328	CACGGTCCAGCTCTTCGCGGAGGACACAGTGGGGAATTTTGGCAATGGGCGCAAGCT	387
O	y	357	GATCCAGCAATGCCCGGTGAGTGAAGAAGGCTTCGGGTTGTAAAGCTCTTTCAGTCGAG	416
D	b	388	GATCAGCGCATGCCCGGTGATGAAGAAGGCTTCGGGTTGTAAAGTACTTTCACGGGG	447
O	y	417	AAGAAAGGTTACGGTAAATATATGTGACTATGACGGTATCGACAAAGAAGCACGGGC	476
D	b	448	AGGAAGGAGTAAGTTAATCTTTCTCATTCACGTTATCCCGCAAGAAGCACGGGC	507
O	y	477	TAACTAGCGTCACGAGCGCGGTAAATACGAGAGGGTCAACGCTTAATCGGAATTAATCG	536
D	b	508	TAACTCGGTGACGAGCGCGGTAAATACGAGAGGGTCAACGCTTAATCGGAATTAATCG	567
O	y	537	GCGTAAAGGTCGCCAGCGCGCTTTGTAAAGTCAGATGTGAATCCCGGGCTTAACCTGG	596
D	b	568	GCGTAAAGCGACCGAGCGGTTTGTAAAGTCAGATGTGAATCCCGGGCTCAACCTGG	627
O	y	597	GAATTCGCTTGAATCTACAGGCTGAGTGGCAGAGGAGGTGGAAATTCATGTGCTA	656
D	b	628	GAATTCGCTGATCTGAGCTGGCAGGCTTGAGTCTCGTAGGGGGGTAGAAATTCAGGTGTA	687
O	y	657	GCAGTGAATCCGTAGAGATATGGAAGAACATCGATGGCGAAGGACGCTCTGGGTAA	716
D	b	688	GCGGTGAATCCGTAGAGATCTGAGAGAAATCCGGTGGCAAGGCGGCGCTTGGACGA	747
O	y	717	CAGTACCGCTATGCGCAAGAACGCGGGAGCAACAGATTAAGTAACTCTGTAGTCCA	776
D	b	748	GACTACCGCTCAGGTGCCAAGACGCGGGAGCAACAGATTAAGTAACTCTGTAGTCCA	807
O	y	777	GCGCCTAAACGATGTCAACTAGTGTGGGCTTATTAAGCTTGG-7TAAGAACTAACG	835
D	b	808	GCGCCTAAACGATGTGCACTTGAGGTGTGCCCTTGAAGGGGTGGCTTCGGAAGCTAACG	867
O	y	836	CGTGAAGTTGACCGGCTGCGGAGATACGGTCCGCAAGTTAAACTCAAGAAATGACGGG	895
D	b	868	CGTGAAGTTGACCGGCTGCGGAGATACGGTCCGCAAGTTAAACTCAAGAAATGACGGG	927
O	y	896	GACCGGCAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAAACTTACCTTA	955
D	b	928	GCGCGGCAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAAAACTTACCTTA	987
O	y	956	CCCTTGAACATGTAGCGAAATTTTCTAGAGAT-AGATTAGTCTTGGGAAAGCTTAACAG	1014
D	b	988	GTCTTGAACATCAACGGAATTTTCTAGAGATGTGCTTCGGGAACCGTGAAGAG	1047
O	y	1015	GTGCTGACATGCTGTGTCACGTCGTCGTCGAGATGTGGTTAAAGTCCGCAACGAGC	1074
D	b	1048	GTGCTGACATGCTGTGTCACGTCGTCGTCGAGATGTGGTTAAAGTCCGCAACGAGC	1107
O	y	1075	GCAACCCCTTATCTTTGTTTCCAGCGGTCCGCGCGGAACTCAAGGAAGATCGCACGTGA	1133
D	b	1108	GCAACCCCTTATCTTTGTTTCCAGCGGTCCGCGCGGAACTCAAGGAAGATCGCACGTGA	1166
O	y	1134	CAAAACCGAGGAAGGTGGGATGACGTCAAGTCTCATGGCCCTTAATGGGATGGGCTTCA	1199
D	b	1168	TAAACTGGAAGGAAGGTGGGATGACGTCAAGTCAATGAGCCCTTACCAACGAGGCTTCA	1222
O	y	1194	CACGTAAATCAATGGCGGTACAGAGGTTGCAACCGCGAAGGGGAGCTAAATTCAGA	1255
D	b	1228	CACGTAAATCAATGGCGGTACAAAGGAAGGACCTCGCGAGAACAGCGGACTTA	1288
O	y	1254	AAGCGGTGTAGTCCGAGTCGGAATCTGCAACTGACTCCGTGAAGTCGGAATGCTGAG	1311

Db 1288 AAGTGGCTAGTCCGAGTTGGACTGCACTGCACTCATGAAGTGGAGCTAG 1347  
Qy 1314 TAATGGCGATCAGCATGTGGGGTGAATACGTTCCCGGGCTTTGTACACACCGCCGTC 1373  
Db 1348 TAATGGTGAATCAGATGTCACGAGTAATACGTTCCCGGGCTTTGTACACACCGCCGTC 1407  
Qy 1374 ACACCATGGAGTGGGTTTCCACAGAACGAGTAACTTAACGTTAAAGAGGCGCTTGCC 1433  
Db 1408 AACACATGGAGTGGGTTTCCACAGAACGAGTAACTTAACGTTAAAGAGGCGCTTGCC 1467  
Qy 1434 ACGGTAGATTCATGACTGGGGTG 1457  
Db 1468 ACTTGTGATTCATGACTGGGGTG 1491  
RESULT 26  
US-11-198-746-158  
; Sequence 158, Application US/11198746  
; Publication No. US20060035256A1  
; GENERAL INFORMATION:  
; APPLICANT: BROW, MARY ANN D.  
; APPLICANT: LYAMICHEV, VICTOR I.  
; APPLICANT: OLIVE, DAVID M.  
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
; TITLE OF INVENTION: PATHOGENS  
; NUMBER OF SEQUENCES: 165  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/11/198,746  
; FILING DATE: 05-Aug-2005  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/941,193  
; FILING DATE: 28-Aug-2001  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: FOPS-01756  
; REFERENCE/DOCKET NUMBER: FOPS-01756  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 158:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1542 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-11-198-746-158  
Query Match 68.3%; Score 995.2; DB 18; Length 1542;  
Best Local Similarity 82.2%; Pred. No. 7.3e-301;  
Matches 1204; Conservative 0; Mismatches 253; Indels 7; Gaps 5;

Qy 1117 GGTAAACGATCGAAAGATGTCCTAATACCGCATATACCTTAAGAGAGAAAGCAGGGATC 176  
Db 148 GATTAATCTAATGAAACGGTAAATACCGCATATACCGTCAACACCAAGAGGGAGACC 207  
Qy 1177 GAAAGACCTTTCGGCTTTTGGAGCGGAGTGTGATTAAGCTAGTTGGTGGGTAAAGGC 236  
Db 208 TTCCGGGCTCTTGGCATTCGATGTCGACAGTGGGATTAAGTATAGTGGGATGACGGC 267  
Qy 237 CTACCAAGCGCAGATCATGATGTTGTTGTGAGAGACGACCAAGCATGAGACTGAGA 296  
Db 268 TCACCTAGGCGACGATCCCTAGCTGTGTCGAGAGATGACCAAGCACAACCTGAGAG 327  
Qy 297 CACGGCCCAAGCTCTCTAAGGAGAGCAGATGGGGAAATTTTGGACATATGGCCGAAGCT 356  
Db 328 CACGGTCCAGCTCTTACGGGAGGACAGATGGGGAAATTTGCAATATGGGCGAAGCT 387  
Qy 357 GATCCAGCAATGCGCGGTGATGTAAGAAAGGCTTCGGGTTGTAAGCTCTTTCAGTCGAG 416  
Db 388 GATGCAAGCATGCGCGCTGTATGTAAGAGGCTTCGGGTTGTAAGCTCTTTCAGCGGG 447  
Qy 417 AAGAAAGTTACGGTAATATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 476  
Db 448 AGAAGAGGAGTAAGTTAATACCTTGTCTCATTTAGCTTACCGCAGAAAGACCGGC 507  
Qy 477 TAATCACTGTCAGACACCGCGGTAAATGTAAGGTCGACCGTTAATCGGAATTAATCG 536  
Db 508 TTAATCGGTGCAACACCGCGGTAAATGTAAGGTCGACCGTTAATCGGAATTAATCG 567  
Qy 537 GCGTAAAGGGTGCAGGCGGCTTGTAAAGTCAATGTAAGTCAATGTAAGTCAATGTAAG 596  
Db 568 GCGTAAAGGGTGCAGGCGGCTTGTAAAGTCAATGTAAGTCAATGTAAGTCAATGTAAG 627  
Qy 597 GAATTCGTTTAAATCAAGAGCTGATGTCAGAGGAGGAGTGAATTCATGTTGTA 656  
Db 628 GAATTCGATCTGATCTGATGTCAGAGGAGGAGTGAATTCATGTTGTA 687  
Qy 657 GCACTGAAATGCTGATGTAATGTAAGTCAATGTAAGTCAATGTAAGTCAATGTAAG 716  
Db 688 GCGGTAAAGGGTGCAGGCGGCTTGTAAAGTCAATGTAAGTCAATGTAAGTCAATGTAAG 747  
Qy 717 CACTGACGCTCATGCAAGAAAGCTGGGGAGCAAGATTAAGTACCTGTAGTCA 776  
Db 748 GACTGACGCTCATGCAAGAAAGCTGGGGAGCAAGATTAAGTACCTGTAGTCA 807  
Qy 777 CGCCCTAAACGATGTAATGTTGTTGGGCTTAATTAAGGCTTGG- TAAAGAGCTAAG 835  
Db 808 CGCCGTAAACGATGTAATGTTGTTGGGCTTAATTAAGGCTTGG- TAAAGAGCTAAG 867  
Qy 836 CGTGAAGTTGACCGCCCTGGGGAGTACGCTGCAAGATTAAGTCAAGAAATTAAGTCAAG 895  
Db 868 CGTGAAGTTGACCGCCCTGGGGAGTACGCTGCAAGATTAAGTCAAGAAATTAAGTCAAG 927  
Qy 896 GACCCGCAAGAGGAGTGAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 955  
Db 928 GACCCGCAAGAGGAGTGAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 987  
Qy 956 CCTTGAATGTAAGCAATTTCTAGAGAT- AGATTAATGTAATGTAATGTAATGTAATGTAATG 1014  
Db 988 GTCTTGAATGTAAGCAATTTCTAGAGAT- AGATTAATGTAATGTAATGTAATGTAATGTAATG 1047  
Qy 1015 GTGCTGATGCTGTCTGCACTGTGTGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1074  
Db 1048 GTGCTGATGCTGTCTGCACTGTGTGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1107  
Qy 1075 GCAACCTTGTCAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1133  
Db 1108 GCAACCTTGTCAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1167  
Qy 1134 CAAACCGAGAAAGGTGGGAGTACGTCAGTCTCAATGAGCCCTTAATGAGGCTTCA 1193  
Db 1168 TAAACTGAGAAAGGTGGGAGTACGTCAGTCTCAATGAGCCCTTAATGAGGCTTCA 1227

Qy	1194	CACGTAATACAAATGGCGGTTACAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAGA	1253
Db	1228	CACGTGCTACAAATGGCGCATACAAAGAGAGCACTTCGCGAGACAAAGCGGACTCATTA	1287
Qy	1254	AAGCGCGTCGTAGTCCGAGTCGGAATCTGCAACTCGACTCCGTGAATCGGATGCTAG	1313
Db	1288	AAGTGCCTGTAGTCCGGAATTGGAACTCTGCAACTCGACTCCATGAAAGTCGGATGCTAG	1347
Qy	1314	TAATTCGGCGGATCAGCATGTCGCGGTGAATPACGTTCCGGGGTTTGTACACACCGCCGCTC	1372
Db	1348	TAATTCGTGGAATCGAATGCGCACGAGTAATACGTTTCCGGGCTTTGTACACACCGCCGCTC	1407
Qy	1374	ACACCATGGAGTGGGTTTCCACGAGACGAGTAATCTTAAACCGTAAGAGGGCGCTTGCC	1433
Db	1408	ACACCATGGAGTGGGTTTGCAAAAGAAAGTAGGTACTTAACCTTCGGGAGGGCGGCTTACC	1467
Qy	1434	AOCGTGAATTCATGACTCGGGGTG	1457
Db	1468	ACTTTGTGATTCATGACTCGGGGTG	1491

RESULT 27  
US-11-198

```

1      ; Sequence 158, Application US/11198794
2      ; Publication No. US20060035257A1
3      ; GENERAL INFORMATION:
4      ; APPLICANT: BROW, MARY ANN D.
5      ; APPLICANT: LYAMICHEV, VICTOR I.
6      ; APPLICANT: OLIVE, DAVID M.
7      ; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
8      ; TITLE OF INVENTION: PATHOGENS
9      ; NUMBER OF SEQUENCES: 165
10     ; CORRESPONDENCE ADDRESS:
11     ; ADDRESSEE: MEDLEN & CARROLL
12     ; STREET: 220 MONTGOMERY STREET, SUITE 2200
13     ; CITY: SAN FRANCISCO
14     ; STATE: CALIFORNIA
15     ; COUNTRY: UNITED STATES OF AMERICA
16     ; ZIP: 94104
17     ; COMPUTER READABLE FORM:
18     ; MEDIUM TYPE: Floppy disk
19     ; COMPUTER: IBM PC compatible
20     ; OPERATING SYSTEM: PC-DOS/MS-DOS
21     ; SOFTWARE: Patent in Release #1.0, Version #1.30
22     ; CURRENT APPLICATION DATA:
23     ; APPLICATION NUMBER: US/11/198,794
24     ; FILING DATE: 05-Aug-2005
25     ; CLASSIFICATION:
26     ; PRIORITY APPLICATION DATA:
27     ; APPLICATION NUMBER: US/09/941,193
28     ; FILING DATE: 28-Aug-2001
29     ; CLASSIFICATION:
30     ; ATTORNEY/AGENT INFORMATION:
31     ; NAME: CARROLL, PETER G.
32     ; REGISTRATION NUMBER: 32,837
33     ; REFERENCE/DOCKET NUMBER: FORS-01756
34     ; TELECOMMUNICATION INFORMATION:
35     ; TELEPHONE: (415) 705-8410
36     ; TELEFAX: (415) 397-8338
37     ; INFORMATION FOR SEQ ID NO: 158:
38     ; SEQUENCE CHARACTERISTICS:
39     ; LENGTH: 1542 base pairs
40     ; TYPE: nucleic acid
41     ; STRANDEDNESS: double
42     ; TOPOLOGY: linear
43     ; MOLECULE TYPE: DNA (genomic)
44     ; US-11-198-794-158

```

Query	1	ATTGAACGGCTGGGGGCGATGCTTTACACATGCAAGTGGAAAGGAGACAGC--GATGCTTGC	58
Query Match	68.3%	Score 995.2; DB 18; Length 1542;	
Best Local Similarity	82.2%	Pred. No. 7.3e-301;	
Matches 1204; Conservative	0;	Mismatches 253; Indels 7; Gaps 5	

Db	28	ATTGAACGCTGGCGGCAAGGCTTAACACATGCAATCGAAACGGTAACAGAAAGAACTTGGC	87
Qy	59	ATC--TGATGGCGAGTGGCGGACCGGAGTGAATATGATCGAAACGTAATCCAAAGAGGGG	116
Db	88	TTCTTTGCTGACGAGTGGCCGGAACGGGTGAATATGTCTGGGAAACCTCGATGATGAAGGGG	147
Qy	117	GGTAACGATCGAAAGATATGCTTAATACCGCATTTACTCTTAAGAGAGAAAGACAGGGATC	176
Db	148	GATAACTACTGGAACCGGTAGCTAATACCGCATTAACCTCGCAACCAAAAGAGGGAGCC	207
Qy	177	GAAAGACCTTGGCTTTTGGAGCGGCGGATGTCATTAAGCTAGTTGGTGGGTAAAGGC	236
Db	208	TTCCGGGCTCTTGGCAATCGAGATGTCGCCAAGATGGGAATTAAGCTAGTAGTGGGTAAACGGC	267
Qy	237	CTACCAAGGCGACATCGATGATTGGTCTGAGAGAGACGACCACTGGAGCTGAGA	296
Db	268	TCACCTTAGGCGACGATCCCTAGCTGGTCTGAGAGATGACCAAGCCACTGGAACCTGAGA	327
Qy	297	CACGGCCGACATCTCTACGGGAGCGACAGTGGGGAAATTTTGGACAAATGGGCGCAAGCT	356
Db	328	CACGGTCCAGACTCCTACGGGAGCGACAGTGGGAAATATGGCAAAATGGGCGCAAGCT	387
Qy	357	GATCCAGCAATGCCCGCTGAGATGAAAGAGGCTTCGGGTTGAAGCTCTTCAGTCCAG	416
Db	388	GATGAGCGCAATGCCCGCTGATGAAAGAGGCTTCGGGTTGAAGTACTTTCAGCGGGG	447
Qy	417	AAGAAAAGGTATACGGTAAATATATGCTGACCTCATGACGGTATCGACAGAAAGACCGGGC	476
Db	448	AGGAAGGAGTAAGTAATTAATCTTTGCTCATTAACGTTAACCCGAGAAAGACCGGGC	507
Qy	477	TAACTACCTGCGCAGCAGCGCGGTAAATCGTAGGGTGCACCGCTTAATCGAAATTAATCTGG	536
Db	508	TAACCTCCGTCGACAGCAGCGCGGTAAATCGAGGGGTGCACCGCTTAATCGAAATTAATCTGG	567
Qy	537	GCGTAAAGGTGCGCAGCGCGCTTTGTAATGACATGTAAGTAATCCCGGGCTTAACCTGG	596
Db	568	GCGTAAACCGCAGCGAGCGGTTGTTAATGATCAAGTGAATATCCCGGGCTCAACCTGG	627
Qy	597	GAATGCGCTTGAACCTACAGAGCTAGAGTGTGGCAGAGGGAGGTGAATTCATATGTGTA	656
Db	628	GAACTGCACTGATATCTGGCAAGCTTGAGTCTCGTAGAGGGGGGTGAAGAAATTCAGAGTGA	687
Qy	657	GCAGTGAATGTGTAAGATATGGAAGAAATCGATGCGCAAGCGCAAGCTCTCTGGGTTAA	716
Db	688	GCGGTGAATGTGTAAGATCTGGAGAAATACCGGTGCGCAAGCGCGCCCTCGACGAA	747
Qy	717	CACGTACGCTCATGCAAGAAACGTTGGGGAGCAAAACAGATTAAGTAACTCCCTGTATGTCGA	776
Db	748	GACTGACCTCTAGGTGCGCAAAACGTTGGGGAGCAAAACAGATTAAGTAACTCCCTGTATGTCGA	807
Qy	777	GCCCTTAAACGATGCAATGATGTTGTTGGGCTTTATATAGGCTTGG--TAACGAAGCTTAACG	835
Db	808	GCCCGGTAAACGATGTCGACTTGGAGGTTGTGCTTGAAGCGGTGGCTTCCGAAGCTTAACG	867
Qy	836	CGTGAAGTGAACCGCTGCGGAGTACCGGTGCGCAAGATTAAACTCAAGAGAAATTGACCGG	895
Db	868	CCTTAAGTGAACCGCTGCGGAGTACCGGCGCAAGTTAAACTCAATGAATTGACCGG	927
Qy	896	GACCCGCAACAGCGGTGATTAATGTGGAATTAATTGATGCAACGGGAAAAAACCCTTAACCTA	955
Db	928	GGCCCGCAACAGCGGTGAGCAATGTGTTTAATTGAATGCAACGGGAAACCTTAACCTG	987
Qy	956	CCCTTGAACATGAAGGAATTTTCTPAAGAT--AGATTAGTCTTTCGGGAACGCTAACAG	1011
Db	988	GTTCTTGAACATCCAGGAAGTTTTCAGAAATGAAGATGTGCTTTCGGGAACGCTGAGACAG	1041
Qy	1015	GTGCTGCATGAGCTGTGTCAGCTCGTGTCTGAGATGTTGGGTAAATGTCCTCGCAACGAGC	1074
Db	1048	GTGCTGCATGAGCTGTGTCAGCTCGTGTGTTGAATGTTGGGTAAATGTCCTCGCAACGAGC	1107
Qy	1075	GCAACCTTGTCAATTAATTCATC--ATTGGTTGGGCACTTAATGAGACTGCGCGGTGA	1133

Db 1108 GCAACCTTATCTTTGTCACAGCGGTCGCGCGGGAACCAAGAGACTGCGCAGTGA 1167  
Qy 1134 CAACCCGGAAGAGTTGGGAGTGAAGTCAAGTCCCTCAAGCCCTTATGGTAAAGGCTTCA 1193  
Db 1168 TAAATGAGAGAGGTGGGATGACGTCAAGTCAATCAAGCCCTTACAGCAAGGCTTCA 1227  
Qy 1194 CACGTAATACATAGCGCGGTACAGAGGGTTCACACCCGCGAGGGAGCTTATCTCAGA 1253  
Db 1228 CACGTCATACATAGCGCATTAAGAGAGGCACTCGCGAGACCAAGCGGACTTCA 1287  
Qy 1254 AAGCGCTGTAGTCCGAGTCCGAGTCTGCAACTCGACTCGGTGAAGTCCGATCTAG 1313  
Db 1288 AAGTCGCTGTAGTCCGAGTCCGAGTCTGCAACTCGACTCGATGAAGTCCGATCTAG 1347  
Qy 1314 TAAATGCGGATCAGATGTCCGCGGTGAATACCTTCCGCGTCTTATACACCCGCGCTC 1373  
Db 1348 TAAATGCGGATCAGATGTCCGCGGTGAATACCTTCCGCGCTTATACACCCGCGCTC 1407  
Qy 1374 ACACCATGGAGTGGGTTTCAACAGAGCAGTGAAGTGAAGAGGCGCTTACC 1433  
Db 1408 ACACCATGGAGTGGGTTTCAACAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1467  
Qy 1434 ACGGTGAGTTCATGACTGGGGTG 1457  
Db 1468 ACTTGATTCATGACTGGGGTG 1491

RESULT 28  
US-11-198-657-158  
; Sequence 158, Application US/11198657  
; Publication No. US2006004029A1  
; GENERAL INFORMATION:  
; APPLICANT: BROW, MARY ANN D.  
; APPLICANT: LYAMICHEV, VICTOR I.  
; APPLICANT: OLIVE, DAVID M.  
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
; TITLE OF INVENTION: PATHOGENS  
; NUMBER OF SEQUENCES: 165  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/11/198.657  
; FILING DATE: 05-Aug-2005  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/941.193  
; FILING DATE: 28-Aug-2001  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: FORS-01756  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 158:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1542 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-11-198-657-158

Query Match 68.3%; Score 995.2; DB 18; Length 1542;  
Best Local Similarity 82.2%; Pred. No. 7.3e-301;  
Matches 1204; Conservative 0; Mismatches 253; Indels 7; Gaps 5;  
Qy 1 ATTGAACGCTGGCGGCAATGCTTTACATGCAAGTCAACGCGACGACG--GATGCTTGC 58  
Db 28 ATTGAACGCTGGCGGCAAGCTTTACATGCAAGTCAACGCGACGCGTAAACGAGAAAGGCTTGC 87  
Qy 59 ATC--TGGTGGCGAGTGGCGGAGCGGTGAGTAATGCAATCGGAACGTATCCAGAAAGAGGGG 116  
Db 88 TTCTTGTGACGAGTGGCGGAGCGGTGAGTAATGCTCGGAACCTCGATGAGAGGG 147  
Qy 117 GGTAAACCATGAAAGATGTCATATACCCCATATACCTTAAGAGGAGAAAGAGCGGATC 176  
Db 148 GATTAATCTAGTGAAGACGTAGCTATACCCCATATACCTCGCAGAACCAAGAGGGGAGC 207  
Qy 177 GAAAGACCTTGGCGCTTTTGAAGCGCGGATGTCATTAAGCTAGTGGTGGGTTAAAGC 236  
Db 208 TTGCGGCTCTTGGCATCGGATGTCGAGATGGGATTAAGTGAAGTGGGTTAAAGC 267  
Qy 237 CTACCAAGCGGACGATGATGTTGTCGAGAGGACGACGACGACATCGGACCTGAGA 296  
Db 268 TCACCTAGGCGGACGATCTTGAAGTGTGAGAGGATGACGACGACGACGACGAACTTGA 327  
Qy 297 CACGCGCCAGACTCTCTACGCGGAGGACGACGAGTGGGAAATTTTGAACAATGGCGCAAGCT 356  
Db 328 CACGCTCCAGACTCTCTACGCGGAGGACGACGAGTGGGAAATTTTGAACAATGGCGCAAGCT 387  
Qy 357 GATTCAGCAATGCGCGTGAAGTGAAGAGCGCTTGGGTTGTAAGCTCTTTCAGTCAAG 416  
Db 388 GATTCAGCAATGCGCGTGAAGTGAAGAGCGCTTGGGTTGTAAGCTCTTTCAGTCAAG 447  
Qy 417 AAGAAAGTTACGCTTAATATATGTCATCATGACGCTATCCAGACAAAGAACGACCGC 476  
Db 448 AAGAAAGTTACGCTTAATATATGTCATCATGACGCTATCCGACAAAGAACGACCGC 507  
Qy 477 TAACTAGTGCAGACGCGGCTATATAGGAGTGAAGGAGTGAAGGAGTGAAGTGAAGTGAAG 536  
Db 508 TAACTAGTGCAGACGCGGCTATATAGGAGTGAAGGAGTGAAGGAGTGAAGTGAAGTGAAG 567  
Qy 537 GCGTAAAGGTTGCGGAGCGGCTTTGTAAGTCAAGTGAATATCCGCGCTTAACTTGC 596  
Db 568 GCGTAAAGGTTGCGGAGCGGCTTTGTAAGTCAAGTGAATATCCGCGCTTAACTTGC 627  
Qy 597 GAATGGCTTTGAAATCTCAAGGCTAGTGTGCGACAGGAGTGAATTTCAATGCTGA 656  
Db 628 GAATGGCTTTGAAATCTCAAGGCTAGTGTGCGACAGGAGTGAATTTCAATGCTGA 687  
Qy 657 GCAGTGAATGCGTGAAGTATGGAAGAACATGATGCGAGGACGCTCTCGGTTAA 716  
Db 688 GCGGTGAATGCGTGAAGTATGGAAGAACATGATGCGAGGAGGAGGAGGAGGAGGAGGAG 747  
Qy 717 CACTGACGCTCATGACGAAAGCGTGGGAGCAAAACGATTAAGTCACTGTGATGCA 776  
Db 748 GACTGACGCTCAAGTGAAGGCGTGGGAGCAAAACGATTAAGTCACTGTGATGCA 807  
Qy 777 GCGCTTAAAGATGTCATCACTGTTGCGCTTATTAAGCTTGG--TAAAGAGCTAACG 835  
Db 808 GCGCTTAAAGATGTCATCACTGTTGCGCTTATTAAGCTTGG--TAAAGAGCTAACG 867  
Qy 836 CGTGAATGACGCGCTGGGAGGATGACGCTGCAAGATTAAGTCAAGGAGTGAAGGAGGAG 895  
Db 868 CGTGAATGACGCGCTGGGAGGATGACGCTGCAAGGATTAAGTCAAGGAGTGAAGGAGGAG 927  
Qy 896 GACCCGCAAGCGGTGATTAATGATTAATTCGATGCAAGCGCAAAACCTTACCTA 955  
Db 928 GACCCGCAAGCGGTGATTAATGATTAATTCGATGCAAGCGCAAAACCTTACCTG 987  
Qy 956 CCCTTGACATGACGAATTTTCTAGAGAT--AGATTAGTCTTGGGAAAGCTTAAACAG 1014  
Db 988 GCTTGACATGACGAATTTTCTAGAGAT--AGATTAGTCTTGGGAAAGCTTAAACAG 1047  
Qy 1015 GTGCTGATGCGTGTCTGACGCTGTGTGTGAGATGTTGGTTAAGTCCGCAAGAGC 1074





Db 251 TAAAGTCGTGATCCGATTTGGAGTCTGCACTCGACTCATGAATGCGAATCGCT 192  
Qy 1312 AGTAATCGCGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTTGTAACACACCGCCG 1371  
Db 191 AGTAATCTTAATCAGAAATGCTACCGGTAAATACGTTCCCGGGCTTTGTAACACACCGCCG 132  
Qy 1372 TCACACCATGGAGAGGGTTTTCACCAAGACAGAGTACTTAACCGTAAAGAGGGCGCTTG 1431  
Db 131 TCACACCATGGAGAGGGTTTTCACCAAGAGTACTTAACCGTAAAGAGGGCGCTTTA 72  
Qy 1432 CCACGCTGAGATTCTCATGACTGGGGTG 1457  
Db 71 CCACCTTGTGATTGATGACTGGGGTG 46  
  
RESULT 30  
US-10-831-286A-48691  
; Sequence 48691, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG  
; APPLICANT: CHATELIER, SONIA  
; APPLICANT: MOIR, DONALD T.  
; APPLICANT: LACROIX, BRUNA  
; APPLICANT: CHILDRESS, DARELL  
; TITLE OR INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
; FILE REFERENCE: 032796-174.001  
; CURRENT APPLICATION NUMBER: US/10/831,286A  
; PRIOR FILING DATE: 2004-04-26  
; PRIOR APPLICATION NUMBER: 60/464,955  
; NUMBER OF SEQ ID NOS: 48788  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 48691  
; LENGTH: 1418  
; TYPE: DNA  
; ORGANISM: Morganella morganii sp. sibirii  
US-10-831-286A-48691  
  
Query Match 67.8%; Score 988; DB 11; Length 1418;  
Best Local Similarity 83.2%; Pred. No. 1.3e-298;  
Matches 1160; Conservative 0; Mismatches 230; Indels 4; Gaps 3;  
  
Qy 68 CGAATGGCGGACGGGTGATGATGATCGGAACGATTCAGAAAGAGGGGGTAAAGCATC 127  
Db 1 CGACCGGGGACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60  
Qy 128 GAAAGATGTCCTAATACCGCATATCTCTAAGAGAGAAAGACAGGGGATCGAAAGACCTTG 187  
Db 61 GAAACGGTACCTAATACCGCATATCTCTAAGAGAGAAAGACAGGGGATCGAAAGACCTTG 120  
Qy 188 CGCTTTTGAGACGGCCGATGTCGATTTAGTGTGAGGTAAAGGCTTACCAAGGCG 247  
Db 121 CACCATCATGATGAACCCATATGGGATTTAGTTGTGATGAGTAAACGCTCACCTTACG 180  
Qy 248 ACGATCACTAGTGTGTCGAAGACGACCAAGCATCTGGGACTGAGACACCGGCCGA 307  
Db 181 ACGATCCCTACTGCTGTGAGAGATGATCAAGCCACATGGGATCGAACAACGGGCCGA 240  
Qy 308 CTCTCAAGGAGGACGACGATGGGAAATTTGACAAATGGGCGCAAGCTTATCCAGCAAT 367  
Db 241 CTCTCAAGGAGGACGACGATGGGAAATTTGACAAATGGGCGCAAGCTTATCCAGCAAT 300  
Qy 368 GCCCGCTGAGTGAAGAGGCTTCGGGTTGTAAAGCTTTCACTCGAAGAAAGGTT 427  
Db 301 GCCCGCTGATGAAGAGGCTTCGGGTTGTAAAGCTTTCACTCGAAGAAAGGTT 360  
Qy 428 ACGTAAATTAATCTGATCTATGACGGTATGACAAAGAACACCGGCTTAATCTGTC 487  
Db 361 AAGGTATATACTCTTGGCAATTTGATGATTCGACGAAAGAACACCGGCTTAATCTGTC 420  
Qy 488 CAGAGCGCGGCTAATACGTAGGGTGAAGGCTTAACTGGAATTAATCTGGGCTAAAGGCT 547  
Db 488 CAGAGCGCGGCTAATACGTAGGGTGAAGGCTTAACTGGAATTAATCTGGGCTAAAGGCT 547

Db 421 CAGACCGCGGTATATACGAGAGGTGCAAGCGTTAATCGAATTAATCTGCGGTAAAGCGC 480  
Qy 548 GCGAGAGCGGCTTTGATAGTACATGTGAATATCCCGGGCTTAATCTGGGAATGGCTT 607  
Db 481 ACGAGAGCGGCTTTGATAGTACATGTGAATATCCCGGGCTTAATCTGGGAATGGCTT 540  
Qy 608 GAACTTAAGAGGTGAGAGTGTGAGAGAGAGTGAATTCATATGTAGCAGTGAATG 667  
Db 541 GATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Qy 668 GATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 727  
Db 601 GATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
Qy 728 ATGCAAGAAAGCGTGGGAGCAACAGATTAATATCCCTGTGTGTCCAGCCCTTAAACG 787  
Db 661 AGGTGCAAAAGCGTGGGAGCAACAGATTAATATCCCTGTGTGTCCAGCCCTTAAACG 720  
Qy 788 ATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 846  
Db 721 ATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
Qy 847 CCGCTGGGAGTACGGTTCGCAAGATTAAATCTCAAGAAATGACCGGGACCGCAAA 906  
Db 781 CCGCTGGGAGTACGGTTCGCAAGATTAAATCTCAAGAAATGACCGGGAGCGGCAAA 840  
Qy 907 GCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 966  
Db 841 GCGGTGAGGATGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 900  
Qy 967 TAGGCAATTTCTGAGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1025  
Db 901 CAGGAACTTAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Qy 1026 CTGTGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085  
Db 961 CTGTGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Qy 1086 CATTAATGCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1143  
Db 1021 CTTTGTGTCAGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
Qy 1144 GAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203  
Db 1081 GAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Qy 1204 AATGCGCTGATCAGAGGTTGCCAACCAGGAGGAGCTTAATCTGAAAGCGCGTGC 1263  
Db 1141 AATGCGCTGATCAGAGGTTGCCAACCAGGAGGAGCTTAATCTGAAAGCGCGTGC 1200  
Qy 1264 TAGTCCGATCGAGTCTGCACTCGATCGATCGATCGATCGATCGATCGATCGATCGAT 1323  
Db 1201 TAGTCCGATCGAGTCTGCACTCGATCGATCGATCGATCGATCGATCGATCGATCGAT 1260  
Qy 1324 TAGCATGTGCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383  
Db 1261 TCAGAAATGCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
Qy 1384 AGTGGTTTTCACCAAGAGAGTATGATTAACGTTAAGAGAGGCGCTTGCACGCTGAT 1443  
Db 1321 AGTGGTTTTCACCAAGAGAGTATGATTAACGTTAAGAGAGGCGCTTACACCTTGTGAT 1380  
Qy 1444 TCATGACTGGGGTG 1457  
Db 1381 TCATGACTGGGGTG 1394  
  
RESULT 31  
US-10-831-286A-48692  
; Sequence 48692, Application US/10831286A  
; Publication No. US20060046246A1  
; GENERAL INFORMATION:  
; APPLICANT: ZENG, QIANDONG



```

APPLICANT: CHATELIER, SONIA
APPLICANT: MOIR, DONALD T.
APPLICANT: LACROIX, BRUNA
APPLICANT: CHILDRESS, DARELL
TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
FILE REFERENCE: 032796-174_001
CURRENT APPLICATION NUMBER: US/10/831,286A
CURRENT FILING DATE: 2004-04-26
PRIOR APPLICATION NUMBER: 60/464,955
PRIOR FILING DATE: 2003-04-24
NUMBER OF SEQ ID NOS: 48786
SOFTWARE: PatentIn version 3.2
SEQ ID NO 48692
LENGTH: 1418
TYPE: DNA
ORGANISM: Morganella morganii sp. sibirici
US-10-831-286A-48692

```

Query Match	67.8%	Score 988;	DB 11;	Length 1418;
Best Local Similarity	83.2%	Pred. No.13e-298;		
Matches 1160; Conservative	0;	Mismatches 230;	Indels 4;	Gaps 3

Oy	68	CGAGGCGCGGACCGSGTGAATATGATCGGAACGTATCCAGAAAGAGGGGGGTAAACGATC	122
Db	1	CGAGCGCGCGACCGSGTGAATATGATGGGGATCTGGCTGATGAGGGGGATTAATCACTG	60
Oy	128	GAAGAATGTGCTAATACCGCATATATCTTAAGAGGAAGACAGGGGATGTAAAGACTTTG	187
Db	61	GAAGCGGTAGCTTAATACGCCATAATATGTCACGGAACCAAGCGGGGACCTTGGCGCTTCG	120
Oy	188	CGCTTTTGGAGCGGCGCGATGTGTGATTTAGTCTAGTTGTGGGGGTAAAGGCTTACCAAGCGG	247
Db	121	CACCACTCAGATGAAACCCATATGGGATTTACCTTGTAGGTGAGGTAAACGAGCTCACCTTAGGGG	180
Oy	248	ACGATCAGATAGTTGGTCTGAGAGGACGACCACTGGGACTGAGACAGCGCCACAGA	307
Db	181	ACGATCTCTTAGCTGTCTTGAGAGATGATCAACCCACACTGGGACTGAGACACGGGCCACAGA	240
Oy	308	CTCTTACCGGAGGACGACAGTGGGGAATTTTGGACAAATGGGCGCAACCTTGATCCAGCAAT	367
Db	241	CTCTTACCGGAGGACGACAGTGGGGAATTTTGGACAAATGGGCGCCCAAGCTTGATCCAGCCAT	300
Oy	368	GCCGCGTGAATGAAGAAGGCCCTTCGGGTGTGAAGCTCTTTCACTGGAAGAAGAAAGATT	427
Db	301	GCCGCGTGTATGAAGAAGGCCCTTCGGGTGTGAAGATCTTTCACTGCGGAGGAAGAGGTGT	360
Oy	428	ACGGTAATTAATCGTGACTCATGACGGTATTCGACAGAAAGAGCACCGGCTTAATCACTGCTC	487
Db	361	AAGGTTAATATACCTTGGCAATTGACGTTACCGACAGAAAGAGCACCGGCTTAATCCGTGTC	420
Oy	488	CAGCAGCCGCGGTAAATCTAGAGGTGCAAGCGTTAATCGAAATTACTGGGCGTTAAAGGAT	547
Db	421	CAGCAGCCGCGGTAAATGCGAAGGATGCAAGCGTTAATCGAAATTACTGGGCGTTAAAGGAT	480
Oy	548	GCGCAGGCGGCTTTGTAATCAGATGTGAATATCCCCGGGCTTAACTGTGGAAATTTGGGTT	607
Db	481	ACGCAAGCGGTTGATTTGAGTCAAGATGTGAATATCCCCGGGCTTAAACCCGGGAATTTGATCT	540
Oy	608	GAAACTACAAAGCTAGATGATGTGGCAAGGAGGATGTGAATTCATGTGTAGCACTGAAATG	667
Db	541	GATACGTGTCAGTTAAGTCTTTGTAAAGGGGGGTAGAAATTCATGTGTAGCGGTGAATG	600
Oy	668	CCTAGAGATATGAGAAACATCGATGGCGAAGGACAGCTTCTGGGTTAACTATGACGCTTC	727
Db	601	CCTAGAGATATGAGAAATATCCGATGGCGAAGGCGGCCCCCTGACAAAGACTGACGCTTC	660
Oy	728	ATGCAAGAAAGCTGGGGAGCAAAACAGATTAATATCCTGTGTATGTCACAGCCCTTAAACG	787
Db	661	AGGTGCGAAAGCTGGGGAGCAAAACAGATTAATATCCTGTGTATGTCACAGCTGTAAACG	720
Oy	788	ATGTCAACTAATGTTGGGCTTAATTAAGCTTGG-TAAGCAACTAAACGGGTAAAGTTGA	846
Db	721	ATGTCAACTTGAAGTTGTGCTTGAAGGCGTGTGCTTCGAGACTAAACGCTTAATGTCGA	780

QY	847	CCGCTGGGGAGTACGGTCGCAAGTTAAATCAAGAATTTGACGGGGACCCGGACAA	906
DB	781	CCGCTGGGGAGTACGGCCGCGCAAGTTAAATCTCAATATTAATTGACGGGGCCCGACAA	840
QY	907	GCGGTGATTTATGTGATTTAATTCGATGCAAGCGAAAACTTACCTTACCTTGCATG	966
DB	841	GCGGTGAGACATGTGGTTTAATTCATGCAAGCGAAGAACTTACCTTACCTTGCATC	900
QY	967	TAGCGAATTTTCTAAGATTAAGTTAGTG-CTTCGGGAACGCTTAACAAGGTGTGATGG	1020
DB	901	CAGAAACTTATAGCAAGAATGCTTGTGTGCTTCGGGAACCTTGAAGACAGGTGTGCAATG	966
QY	1026	CTGTGTCAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGT	1088
DB	961	CTGTGTCAGCTCGTGTGCTTGTGAATGTTGGTTAAGTCCCGCAACGAGCGCAACCTTAT	1020
QY	1086	CATTATATTCGCATCATTTTG--GTTGGGCACTTTAATGAACTGCGGTGACAAACCGGAG	114
DB	1021	CGTTTGTGGCCAGCCCGTATGCGCGGAACCTCAAGGAACATGCGGTGATTAACCGGAG	108
QY	1144	CAAGGTGGGGAGAGAGTCGAATGCTCATGAGCCCTTATGGGTAGGGGCTTCAACGTAATAC	120
DB	1081	GAAAGTGGGAGTAGAGTCGAATGATCATGTGGCCCTTACGATAGGGCTACACAAGTGTAC	114
QY	1204	AATGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAACGCGCTCG	12
DB	1141	AATGCGGTATACAAAGGGAAGCGAACCTTCGAAGCAACGGAACCTAATTAATGACTCG	12
QY	1264	TAGTCCGATCCGAGTCTTGCAACTGCACTCCGTAAGTGGATTCGTATTAATCCGGGA	13
DB	1201	TAGTCCGATTTGAGTCTTGCAACTGCACTCCGTAAGTGGATTCGTATTAATCCGTAAG	12
QY	1324	TGAGCATGTGCGGGTGAATACGTTCCCGGAGCTTGTACACACCGCCGTCACACATGGG	136
DB	1261	TGAGATGTCTAGGTGAATACGTTCCCGGAGCTTGTACACACCGCCGTCACACATGGG	13
QY	1384	AGTGGATTTCACACAGAACGTAAGTCTAACCGTAAAGAGGGGCGCTTGCCACGATGAGT	14
DB	1321	AGTGGATTTCACAAAGAAATGATGACTTAACCTTCGGGAGGGGCGCTTAACCACTTTGTAT	13
QY	1444	TCATGACTGGGGTG 1457	
DB	1381	TCATGACTGGGGTG 1394	

## RESULT 32

```

US-10-831-286A-48667/c
; Sequence 48667, Application US/10831286A
; Publication No. US20060046246A1
GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831, 286A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464, 955
; PRIOR FILING DATE: 2003-04-24
; NUMBER OF SEQ ID NOS: 48788
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48667
; LENGTH: 1436
; TYPE: DNA
; ORGANISM: Cedecaea lapagei
US-10-831-286A-48667

```

Query Match	67.3%	Score 981.2;	DB 11;	Length 1436;
Best Local Similarity	82.8%	Pred. No. 1.7e-296;		
Matches 1157; Conservative	0;	Mismatches 238;	Indels 3;	Gaps 3

```
OY      63  GGTGGCGAGTGGCGGACGGGTGATGATGCATCGGAAGTATCCAGAAAGAGGGGGGTAA 122
Db      1431 GGTGACGAGCGGCGGACGGGTGATGATGTCTGGGATCTGCTTATGAGAGGGGATTAAC 1372
OY      123  GCATCGAAAGATGTGCTAATAACCGCATATCTTAAGAGAGAAAGACAGGGGATCGAAAGA 182
Db      1371 TACTGAAACGGTAGTACTAATACCGCATACGTGCAAGACCAAAAGAGGGGGACTTTCGGG 1312
OY      183  CTTTCCGCTTTTGGAGCGGCGCATGTCTGATTAAGTAACTGTTGGGTAAAGGCTTACCA 242
Db      1311 CCTTTGCGATCAATGAAACCGAATGGGATTTAGTAGTGGGTGAAACGGCTACACT 1252
OY      243  AGGCGACGATCATGATGTTGGTCTGAGAGGACGACGACGACCTGGGACTGAGACAGGAC 302
Db      1251 AGGCGACGATCTCTAGTGTGCTGAGAGGATGACAGGACCACTGGAACTGAGACAGGT 1192
OY      303  CCAGACTCTTACGGGAGGACAGCATGTGGGAAATTTTGAACAATGGGCGCAAGCTGATCCA 362
Db      1191 CCAAGCTCTACGGGAGGACAGCATGTGGGAAATTTTGAACAATGGGCGCAAGCTGATCCA 1132
OY      363  GCATGCGCGGTGATGAAAGAGCCCTTGGGTTGTAAGCTCTTCACTGAGAGAA 422
Db      1131 GCCATGCGCGGTGTATGAAAGAGCCCTTGGGTTGTAAGCTTTCAGCGAGAGGAAG 1072
OY      423  AGGTTAGGTTAAATTAATGTGATCATGACGGTATCGACAGAAAGACCGGCTAATCTA 482
Db      1071 GCGATTAAGTTAATACTTGTTCATTAACGTTACTTCCGAGAAAGACCGGCTAATCTC 1012
OY      483  CGTGCACAGACCGCGGTATATACGTAGGGGTGCAAGCGTTAATCGAATTAATCTGGCGTAA 542
Db      1011 CGTGCACAGACCGCGGTATATACGTAGGGGTGCAAGCGTTAATCGAATTAATCTGGCGTAA 952
OY      543  AGGTCGCCAGCGGCTTTGTAAGTCAATGTGAAATCCCGGGCTTAACTGGGAATG 602
Db      951 AGCGACCGCAGCGGCTTTGTAAGTCAATGTGAAATCCCGGGCTTAACTGGGAATG 892
OY      603  CGTTGAAACTCAAGGCTAGAGTGTGCAAGGAGGTGAAATTCATGTGTAGCACTG 662
Db      891 CATTGAAACTCGGCAAGCTTAGTCTTTAAGAGGGGGTAAATTCAGAGTGTAGCGGTG 832
OY      663  AAATGCGTAGATATGGAAGAACATCGATGCGCAAGGACCTTCTGGGTTAACTGTA 722
Db      831 AAATGCGTAGATATGGAAGAACATCGATGCGCAAGGACCTTCTGGGTTAACTGTA 772
OY      723  CGCTCATGCAAGAAAGCGTGGGAGCAAAACAGAAATTAATTAATCCTTGTAATGTCACGCTT 782
Db      771 CGCTCATGCTGCAAAAGCGTGGGAGCAAAACAGAAATTAATTAATCCTTGTAATGTCACGCTT 712
OY      783  AAAGGATGTCAACTAGTTGTTGGGCTTAATTAAGCTTGG-7AAAGGATGTCAAAGCGTGA 841
Db      711 AAAGGATGTCAACTAGTTGTTGGGCTTGGAGGTGGCTTCCGAGCTTAACGGGTAA 652
OY      842  GTTGACCGCTTGGGAGTACGGTGCAGAAATTAATACTCAAAAGAAATTGACGGGAGCCG 901
Db      651 GTTGACCGCTTGGGAGTACGGTGCAGAAATTAATACTCAAAAGAAATTGACGGGAGCCG 592
OY      902  CACAAGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAACCTTACTCTCTTG 961
Db      591 CACAAGCGGTGATTAATGTGATTAATTCGATGCAACCGGAAACCTTACTCTCTTG 532
OY      962  ACATGATGCGGAATTTCTAGAGATGATTAATG-CTTGGGAAACGCTTAACAGGTGTG 1020
Db      531 ACATGATGCGGAATTTCTAGAGATGATTAATG-CTTGGGAAACGCTTAACAGGTGTG 472
OY      1021  CATGGCTGTGCTAGCTGTGTGAGATGTTGGGTTAAATCCGCAACGAGCGCAAC 1080
Db      471 CATGGCTGTGCTAGCTGTGTGAGATGTTGGGTTAAATCCGCAACGAGCGCAAC 412
OY      1081  CTTGTCAATTAATGCGATC-ATTGGTTGGGCACTTAAATGAGACTGCGGTGACAAAC 1139
Db      411 CTTATCTTTGTTGGCAGCGGTTGGCGGGAACCTCAAAAGAGACTGCGCAATGTAATAC 352
```

```
OY      1140 GAGGAAGTGGGGATGACGTCAAGTCTCTATGCGCCCTTAATGGGTAGGGCTTCAACGTA 1199
Db      351 GAGGAAGTGGGGATGACGTCAAGTCTCTATGCGCCCTTAATGGGTAGGGCTTCAACGTA 292
OY      1200 ATCAATAGCGCGGTACAGAGGCTTCCCAACCCCGAGGGGAGCTTAATTCAGAAAGCC 1259
Db      291 CTCAATAGCGCGGTACAGAAAGAGAGACCTTCGGAAGAGCAAGCGGACCTCATTAAGTGC 232
OY      1260 GTGCTAGCGGGATCGAGTGTGCACTCGACTCGTGAAGTGGGAATCGCTAATATCG 1319
Db      231 GTGCTAGCGGGATCGAGTGTGCACTCGACTCGTGAAGTGGGAATCGCTAATATCG 172
OY      1320 CGGATCAGCATGTGCGCGGTAAATAGTTCGCGGCTCTTGTACACACCGCCGTCACCA 1379
Db      171 TAGATCAAGATGCTACAGGTGAATAGTTCGCGGCTCTTGTACACACCGCCGTCACCA 112
OY      1380 TGGGATGGGTTTACCAAGAACAGTATGCTTAACGTTAAGAGAGGCGCTTGCCAGGTG 1439
Db      111 TGGGATGGGTTTCAAAAGAAAGTATGATGTTAATCTTCGGAGAGGCGCTTACCACTTTG 52
OY      1440 AGATTCATGACTGGGGTG 1457
Db      51 TGATTCATGACTGGGGTG 34
```

```
RESULT 33
US-10-831-286A-48672
; Sequence 48672, Application US/10831286A
; Publication No. US20060046246A1
; GENERAL INFORMATION:
; APPLICANT: ZENG, QIANDONG
; APPLICANT: CHATELIER, SONIA
; APPLICANT: MOIR, DONALD T.
; APPLICANT: LACROIX, BRUNA
; APPLICANT: CHILDRESS, DARRELL
; TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES
; FILE REFERENCE: 032796-174.001
; CURRENT APPLICATION NUMBER: US/10/831,286A
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/464,955
; NUMBER OF SEQ ID NOS: 4878
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48672
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Citrobacter Youngae
US-10-831-286A-48672
```

```
Query Match      67.2%; Score 979.8; DB 11; Length 1481;
Best Local Similarity 81.7%; Pred. No. 4; Be-296;
Matches 1181; Conservative 0; Mismatches 257; Indels 7; Gaps 4;

OY      20  CTTTACATGCAATGCAACGCGACACGAGT-CTTGATCTGTGGCGAGTGC 75
Db      1  CTTAACAATGCAAGTCAAGCGTGAACAGAGAGCTTGCTCTTGGTGAACGAGTGC 60
OY      76  GAGCGGTGATGATGATGCAATGGAACGTATCCAGAAAGGGGGTAAACGCAATCGAAATG 135
Db      61  GAGCGGTGATGATGATGCTGGGAAACTGCGATGAGGGGGATATACTTGTGAAAGGT 120
OY      136  TGTCAATACCGCATATCTCTAAGGAGAAAGCAGGGGATCGAAAGCTTGCGCTTTTG 195
Db      121  AGCTAATATCGCATATCGTGCAGAACCAAGAGGGGAACTTTCGGGCTCTTGCCATCG 180
OY      196  GAGCGCGCATGATGCTGATTAAGTATGTTGGGTAAAGGCTTAACAAAGCGAGATCAG 255
Db      181  GATGTGCGCAGATGAGATTAAGTATGATGATGAGGTAAACGCTCAACAGGAGAGATCCC 240
OY      256  TAGTGTCTGAAGAGACGACCGCACTGGGACTGGAACACGCGCCGATCTTACG 315
Db      241  TAGTGTCTGAAGAGATGACCGCACCACTGGAACTGGAACACGCGTCCAGATCTTACG 300
```

316 GGAGCAGCAGTGGGGAATTTTGA CAA TGGGCGCAGCCTGATCCAGCAATGCCGCTG 375  
301 GGAGCAGCAGTGGGGAATTTGCA CAA TGGGCGCAGCCTGATCCAGCAATGCCGCTG 360  
376 AGTGAAGAAGCCTTGGGTTGTAAAGCTCTTTGCTGAGAAAGAAAAGTTACGTTAAA 435  
361 TATGAAGAAGCCTTGGGTTGTAAAGCTCTTTGCTGAGAAAGAAAAGTTACGTTAAA 420  
436 TAATGTCATCATGACGGTATCGA CAG AAGAAAGACCGGCTTAATCTAGTCCAGCAGCC 495  
421 TAACTTGAAGCAATTAACCTTACT CCG AAGAGACACCGGCTTAATCTAGTCCAGCAGCC 480  
496 GCGGTAATACGTAGGGGTCAAGCCTTAA TCGGAATTAATCGGCGTAAAGGTCGCGAGCC 555  
481 GCGGTAATACGTAGGGGTCAAGCCTTAA TCGGAATTAATCGGCGTAAAGGTCGCGAGCC 540  
556 GCGTTGTAAAGTCAAGTGTAAATCCCGGGCTTAACTTGGGAATTCGCTTTGAAACTAC 615  
541 GGTCTGTCAAGTCGATGTGAATCCCGGGCTTAACTTGGGAATTCGCTTTGAAACTAC 600  
616 AAGGTAAGTGTGCGACAGAGGAGTGAATTC AATGTGTAGCAGTGAATTCGTAAGAG 675  
601 CAGGTAAGTGTGCGACAGAGGAGTGAATTC AATGTGTAGCAGTGAATTCGTAAGAG 660  
676 TATGGAAGAACATCGATGCGAAGCGAGCCTCTGGGTTTAACTGACGCTCATGACGCA 735  
661 TCTGAGAGAAATCCGCTGCGAAGCGAGCCTCTGGGTTTAACTGACGCTCATGACGCA 720  
736 AAGCGTGGGAGCAAA CAG AATTAAGATACCTGTGTAGTCAAGCCTTAAAGATGCAAC 795  
721 AAGCGTGGGAGCAAA CAG AATTAAGATACCTGTGTAGTCAAGCCTTAAAGATGCAAC 780  
786 TAGTGTGGGCTTATTAGGCTTG- TAA CCAAGCTAACGGGTGAAGTGAACGCTG 854  
781 TTGAGAGTTGTGCTTGAAGCGTGTGCTTCCGAGCTTAACGGGTTAAGTGAACGCTG 840  
855 GGAAGTACGGTCCGAAGATTAATACTCA AAGAAATGACCGGGGACCCGCAAGACGGGTGA 914  
841 GGAAGTACGGTCCGAAGATTAATACTCA AAGAAATGACCGGGGACCCGCAAGACGGGTGA 900  
915 TTATGTGATTAATTTGATGCAACGCGAAA AACTTAACTTACCTTGAATGAGGAT 974  
901 GATGTGTGTAAATTTGATGCAACGCGAAA AACTTAACTTACCTTGAATGAGGAT 960  
975 TTTCTGAAGATTAATTAAGT-CTTGGGGAACGCTAACAGAGTGCATGCTGCTGCTC 1033  
961 TTTCTGAAGATTAATTAAGT-CTTGGGGAACGCTAACAGAGTGCATGCTGCTGCTC 1020  
1034 AGCTGTCGTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAACCTTGTCAATTAAT 1093  
1021 AGCTGTCGTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAACCTTGTCAATTAAT 1080  
1094 GCCATCA-TTTGGTTGGGCACTTTAATGAATGCTGCGGTGA CAAACCGAGAGAGGTGG 1152  
1081 GCCAGCGGTTCCGCGCGGAACCTCA AAGAGACTGCCAGTGAATAACTGAGAGAGGTGG 1140  
1153 GATGAGTCAAGTCTCTAAGGCGCTTAA TGGGTAGGCTTCAACGTAATTAATGAGCGG 1212  
1141 GATGAGTCAAGTCTCTAAGGCGCTTAA TGGGTAGGCTTCAACGTAATTAATGAGCGG 1200  
1213 TACAGAGGTTGCGCAACCGCGAGGGGAGCTAATCTCAGAAAAGCGGTGTAAGTCCGGA 1272  
1201 TACAAAGAGAGGAGCCTCGCGAGAGCAAGCGA CTTCAATTAAGTGTGTAGTCCGGA 1260  
1273 TCGAGATCTGCAACTCGACTCGGTGAAGTGAATGCTAATGTAATGCGGATCAAGATG 1332  
1261 TTGGAAGTCTGCAACTCGACTCGGTGAAGTGAATGCTAATGTAATGCGGATCAAGATG 1320  
1333 GCGGCGTAATACGTTCCCGGCTTGTGA CACACCGCGCTCAACCAATGGAAGTGGGTT 1392  
1321 CACGCGTAATACGTTCCCGGCTTGTGA CACACCGCGCTCAACCAATGGAAGTGGGTT 1380  
1393 CACCAAGAGAGGATGTAACGTAAGAGAGGCGCTTGCACAGGTGAGATTCATGACTG 1452

DB 1381 CAAAAGAGTAAAGTAACTTAACCTTGGAGAGGCGCTTACCACTTTGTGATTCATGACTG 1440  
QY 1453 GGGTG 1457  
DB 1441 GGGTG 1445

RESULT 34  
US-10-831-286A-48689/c  
Sequence 48689, Application US/10831286A  
Publication No. US20060046246A1  
GENERAL INFORMATION:  
APPLICANT: ZENG, QIANDONG  
APPLICANT: CHATELIER, SONIA  
APPLICANT: MOIR, DONALD T.  
APPLICANT: LACROIX, BRUNA  
APPLICANT: CHILDRESS, DARRELL  
TITLE OF INVENTION: GENUS, GROUP, SPECIES AND/OR STRAIN SPECIFIC 16S rDNA SEQUENCES  
FILE REFERENCE: 032796-174.001  
CURRENT APPLICATION NUMBER: US/10/831, 286A  
CURRENT FILING DATE: 2004-04-26  
PRIOR APPLICATION NUMBER: 60/464,955  
PRIOR FILING DATE: 2003-04-24  
NUMBER OF SEQ ID NOS: 48788  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 48689  
LENGTH: 1525  
TYPE: DNA  
ORGANISM: Enterobacter hormaechei  
US-10-831-286A-48689

Query Match 67.1%; Score 977.8; DB 11; Length 1525;  
Best Local Similarity 82.4%; Pred. No. 2e-295;  
Matches 1194; Conservative 0; Mismatches 247; Indels 8; Gaps 6;

QY 17 ATGCTTACATGCAAGTCCAGACGACG--GATGCTTGACATCT--GGTGGGAGCT 72  
DB 1510 AGGCTTAACATGCAAGTCCAGACGCTTAACAGAAAGCAGCTTGCTGCTTCCGTGACAGCT 1451  
QY 73 GCGGAGCGGAGTAAATGATCGAATCGGAACGTATCCAGAAAGGAGGAGTAAACGATCGAAAG 132  
DB 1450 GCGGAGCGGAGTAAATGATCGAATCGGAACGTATCCAGAAAGGAGGAGTAAACGATCGAAAG 1391  
QY 133 ATGTCTAATACCGCATTAATCTTAAGAGAGAAACGAGGAGTCAAAAGACCTTGCGCTT 192  
DB 1390 GGTAGCTAATACCGCATTAATCTTAAGAGAGAAACGAGGAGTCAAAAGACCTTGCGCTT 1331  
QY 193 TTGAGACGCGCATGTCTGATTAAGTGAATGTTGGTGGGTAAAGCCTTACCAAGCGACGAT 252  
DB 1330 TCGAGTGTGCCAGATGGGATTAAGTGAATGTTGGTGGGTAAAGCCTTACCAAGCGACGAT 1271  
QY 253 CAGTATGTTGCTGAGAGAGCAGACGACGACACTGGGACTGAGACAGCGCCGACACTCCT 312  
DB 1270 CCTAGTGTGCTGAGAGAGATACAGCCACTGGAATCTGAGACAGCGTCCAGACTCCT 1211  
QY 313 ACAGGAGCAGCAGTGGGGAATTTTGA CAA TGGGCGCAGCCTGATCCAGCAATGCCGCTG 372  
DB 1210 ACAGGAGCAGCAGTGGGGAATTTTGA CAA TGGGCGCAGCCTGATCCAGCAATGCCGCTG 1151  
QY 373 GTGAGTGAAGAAGCCTTGGGTTGTAAAGCTCTTTCAATCGAAGAAAGAAAGTTACGTT 432  
DB 1150 GTGATGAAGAAGCCTTGGGTTGTAAAGCTCTTTCAATCGAAGAAAGAAAGTTACGTT 1091  
QY 433 AATTAATCTGATCTCATGACGGTATCGA CAG AAGAAAGACCGGCTTAATCTAGTCCAGCA 492  
DB 1090 TAATAATCTCAAGATTTGACGTTACCTCGCAAGAAAGACCGGCTTAATCTAGTCCAGCA 1031  
QY 493 GCGCGGTAAATACGTAGGGTGAAGCGTTAATCGAATTAATGAGGCTAAAGGAGTGCAG 552  
DB 1030 GCGCGGTAAATACGTAGGGTGAAGCGTTAATCGAATTAATGAGGCTAAAGGAGTGCAG 971  
QY 553 GCGCGCTTTGTAAATGATGTAATTCGCCGGGCTTAACTTGGGAATTCGCGTTTGAAC 612

Db	970	GGCGGCTCTGTCAAGTCGATGATGTGAATCCCGGGCTCAACTGGGAATCTGCATTGGAAC	911
Qy	613	TACAAAGGCTAAGATGTGGCAGAGGGAGTGAATTCATGTGTAGCAGTGAATGCGTAG	672
Db	910	TGGCAGGGCTAGAGTCTTGTAGAGGGGGGTGAAGATTCAAGTGTAGCGGTGAATGCGTAG	851
Qy	673	AGATATGGAAGAATCATCGATGGCGAAGGACGCTCTGGGTTAACTGAACGCTCATGCA	732
Db	850	AGATCTGGAGGAATACCGGTGGCGAAGGGGGCCCCCTGGACAAGAATGACCTCAGGTG	791
Qy	733	CGAAAGCGTGGGAGCAACAGAAATTAGTACCTGTGTGTGCCAGCCCTTAACAGATGTC	792
Db	790	CGAAAGCGTGGGAGCAACAGAAATTAGTACCTGTGTGTGCCAGCCCTTAACAGATGTC	731
Qy	793	AACTAGTGTGGGCTTATTTAGGCTTGG-TAACGAAGCTAACCGGTGAAGTTGACCGGC	851
Db	730	GACCTTGGAGGTGTGTCCCTTGAGGCGTGGCTTCCGAGCTAACGCGTTAATGTCGACCGCC	671
Qy	852	TGGGGAGTACCGGTCCGCAAGATTAAATCTCAAGAAATTTGACGGGGACCCGCAACGCGT	911
Db	670	TGGGGAGTACCGGTCCGCAAGATTAAATCTCAAGAAATTTGACGGGGACCCGCAACGCGT	611
Qy	912	GGATTATGTGATTTAATTCGATGACACGCGAAAAACCTTACCTACCTTGACATGTAGCG	971
Db	610	GGAGCATGTGGTTATTATTCGATGACACGCGAAAGAACCTTACTTACTTTGACATCGAGG	551
Qy	972	AATTTTCTAGAGATAGATTAGTGTG-CTTGGGGAAACGTTACACAGGTGCTGATGGCTGTC	1033
Db	550	AACTTAGCAGAAATCTTGTGTGCTTGGGAACTCTGAGACAGGTGCTGACATGGCTGTC	491
Qy	1031	GTCACTCGTGTGCTGAGATGTTGGGTTAATGTCGCGCAACGAGCGCAACCCCTGTCAATTA	1090
Db	490	GTCACTCGTGTGTTGGAATGTTGGGTTAATGTCGCGCAACGAGCGCAACCCCTTATCTTT	431
Qy	1091	ATTGCGCATCA-TTTGTTGGGCACTTTAATGAGACTGCCGTGACAAACCGGAGGAAGGT	1145
Db	430	GTTCGCACGCGTTTAGCGCGGGAATCAAGAGAGACTGCCAGTGAATTAACCTGGAGGAAGGT	371
Qy	1150	GGGATGACGTCAATCTCTCATAGGCCCTTATATGGGTATAGGCTTCAACAGTAATACAAATGGC	1205
Db	370	GGGGATGACGTCAATCTCTCATAGGCCCTTATCAAGTAGGGGCTTACACACGCTGTCACATGAGC	311
Qy	1210	GGGTACAAAGGGTGTCCCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGCGTGTAGTCC	1265
Db	310	GCATACAAAGAGAGACGCACTCTGCCAGAGCAAGCGGACCTCAATAAGTGTGCTGTAGTCC	251
Qy	1270	GGATCGAGTCTGCAACTCGACTCCGTAAGTGGGAATTCGTAATTCGCGGATCAGCA	1325
Db	250	GGATTGGAGTCTGCAACTCGACTCCATGAAGTCGGAATTCGTAATTCGCGGATCAAGA	191
Qy	1330	TGTGGCGGTGAATACGTTCCGCGGTCTTGTATACACCGCCGCTACACATAGGGAGTGG	1385
Db	190	TGCCACGGGTGAATACGTTCCGCGGTCTTGTATACACCGCCGCTACACATAGGGAGTGG	131
Qy	1390	TTTCAACCAAGACAGTAGTCTTAACCTGTAAAGAGGGGCGCTT-GCCACGATGAGATTCAATG	1448
Db	130	TTGCAAAAGAGTAGTAGTCTTAACCTTCCGGAGGGCGCTTAACCACTTTGTGATTCATG	71
Qy	1449	ACTGGGGGTG 1457	
Db	70	ACTGGGGGTG 62	

RESULT 35  
US-11-055-637-76

```

; Sequence 76, Application US/11055637
; Publication No. US20050260619A1
; GENERAL INFORMATION
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: DUBOIS, Jason
; APPLICANT: EDGE, Tom
; APPLICANT: MASSON, Luc

```

```

; APPLICANT: TREVORS, Jack T
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
; TITLE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES
; FILE REFERENCE: 2139-330S
; CURRENT APPLICATION NUMBER: US/11/055,637
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,288
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Probe for DNA array
US-11-055-637-76

```

Query Match	58.4%;	Score 850.4;	DB 17;	Length 1485;
Best Local Similarity	76.1%;	Pred. No. 1.6e-255;		
Matches 118;	Conservative	0;	Mismatches 336;	Indels 16; Gaps 5

Oy	4	GAACGCTGGCGGCAATCCTTTAAACATTCGAAGTCGGAACGGCAGACCGATGCTTGCATCTG	63
Db	4	GAACGCTGGCGGCAATCCTTTAAACATTCGAAGTCGGAAGTCGGAAGCTTCTCTT	63
Oy	64	GTGCGCAGTGGCGGACGGGTGAGTATGATCGG--ACGTATCCAGAAAGAGGGGGGTAACT	122
Db	64	GATTTTAGGGGGGAGACGGGTGAGTAAACGTGGGCAACTGCTCTACAGATGGGGATTAC	122
Oy	123	GCATCGAAGAATGTGCTAATCCGCATATCTTAAGAGAAAGCAGGGGATCGAAA--	180
Db	124	TCCGGGAAACCGGGGCTAATACCGAATATACGTTTGTCCCATGACAACTCTGAAAG	183
Oy	181	-----GACCTTGGCCTTTTGGACGGCCCATGCTGATTAGCTAGTGGTGGGGTAA	233
Db	184	ACGGTTTCGACTGTCACTGTAGAAAGGGCCCGCGCATTAGCTAGTGGTGGGTAAAT	243
Oy	234	GGCCTACCAAGGCGACGATCAGTATTGGTCTGAGAGGACGACAGCCACACTGGGACTG	293
Db	244	GGCCTACCAAGGCGACGATGCTTACCACCTGAGAGGGTATCGGCCACACTGGGACTG	303
Oy	294	AGACACGGCCCAAGACTCTTACGGGAGGACGACAGTGGGAAATTTTGACCAATGGCGCAAG	353
Db	304	AGACACGGCCCAAGACTCTTACGGGAGGACGACAGTGGGAAATCTTCCACCAATGACGAAG	363
Oy	354	CCTGATCCAGCAATGCCCGGTGATGTAAGAAAGCCTTGGGTGTAAAGCTTTCAATC	413
Db	364	TCTGATGAGCAATGCCCGGTGAGGAAAGGTTTTCGATCGTAAAGCTCTGTGTAA	423
Oy	414	GAGAGAAAGGTTACCGTTAAATATCGTGACTCATGACGCTATTCGACAGAAAGACACC	473
Db	424	GGAGAAACAAGTACGGGAGTAACTGCCCCGTGCCATACGCTTACTTTTGAAGAGCCAC	483
Oy	474	GGCTAACTACGTGCCAGACGCCGGGTAAATCGTAGGCTGCAGCGTTAATCGGAATTAC	533
Db	484	GGCTAACTACGTGCCAGACGCCGGGTAAATCGTAGGCTGCAGCGTTAATCGGAATTAT	543
Oy	534	TGGGCTAAAGGGGTGCGACAGCGGCTTTGTGTAATGCATGTGAATATCCCGCGCTTAAC	593
Db	544	TGGGCTAAAGGGCGCGACAGCGGCTTTGTGTAATGTGTAAGTGAAGGCCACCGCTCAAC	603
Oy	594	TGGGAATTCGTTTGAACAACAGGCTAGAGTGGCAGAGGGAGGTGGAATTCATGT	653
Db	604	GTGAGAGGCTATTGGAAACTGGAACAATTGAGTACAGAAAGGAAACGGAAATTCACGT	663
Oy	654	GTAGCAGTGAATTCGCTAGAGATATGAAAGACATCGATGGCGAAGGCAAGCTCTGGGT	713
Db	664	GTAGCGGTGAATTCGCTAGAGATGTGGAAGAACACAGTGGCGAAGGCGGCTTCTGGTC	723
Oy	714	TAAACATGACGCTTACATGACAGAAAGCTGTGGGAGCAACAGGATTAATATCCTCTGGTAT	773
Db	724	TGTAACTGACGCTGAGGGCGCGAAACCGTGGGAGCAACAGGATTAATATCCTCTGGTAT	783

```

Oy      774   CCAGCGCCCTMAAACGATGTCAACTTAATTGTT--GGGCCTTAAATAGCCTTGTAACGAAGCT    831
Db      784   CCACGCCCTTAAACGATGAGTGTCTAAGTGTTAGGGGGTTTTCCGCCCTTAACTGCTGCAGCT    843
Oy      832   AACGGGTGAAGTTGACCCGCTTGAGAGTACGGTCGCAAGATTAAAACTCAAAGAAATTGA    891
Db      844   AAGCATTTAAGACATCTCCGCTTGAGAGTACGGTCGCAAGATTAAAACTCAAAGAAATTGA    903
Oy      892   CCGGGACCCTGCACAAGCCGGTGTATTATGTGATTAATTGCATGCAACGCCAAAACCTTA    951
Db      904   CGGGAGCCCGCACAAAGCGGTGAGCATGTGTGTTAAATTGCMAACCAACGCTAAAGAACTTA    963
Oy      952   CCTACCCCTTGAATGATGAGGAA---TTTCTAGAGTAGATTAATGATGTGCTTGAGAACGCTA    1008
Db      964   CCAGGCTTTTGACATCCCACCTGACCCGAGTGAAGATAACGCTTTCCTTCCTTGCGGGAGCAAGTGG    1023
Oy      1009   ACACAGGTCTGCATGCTGTCGTCACTCGTGTGAGATGATGAGATTTGGGTTAAAGTCCGCA    1068
Db      1024   TGAACAGTGTGATGATGTTGTGTGTGTCAGCTGTGTGCGAAGATGATGGGTTAAAGTCCGCA    1083
Oy      1069   ACGAGCGCAACCTTGTCTAATTAATTGCCATCATTTGTGTGGGCACTTTAATAGAATGCC    1128
Db      1084   ACGAGCGCAACCTTGTCTAATTAATTGTTCGACATTCAGTTGGGCACTTAAAGTGAATGCC    1143
Oy      1129   GGTGACAAACCGGAGGAAGGTGGGGAATGACGTCAAGTCTCATGTGACCTTATGCGGTAAGG    1188
Db      1144   GGATGATTAACCGGAGGAAGGTGGGGAATGACGTCAATTCATCATCCCCCTTATGACCTGGG    1203
Oy      1189   CTTCACACGTTAATATACATGCGCGTACAGAGGTTGCAACCCGCGAGGGGGAGCTAATC    1248
Db      1204   CTACACACGTGTCAATATGATGATATACAGAGGTTGCAACCCGCGAGGGGGAGCCCAATC    1263
Oy      1249   TCAGAAAGCGCGCTGTGATGTCGAGTCGGAATCTTGCAACTGCACTCCGCTGAAGTGGGAATC    1308
Db      1264   CCATAAATTCGTTCCAGTTCGATGGAAGCTGCAACTCGCTTCACATGAAGTGGGAATC    1323
Oy      1309   GCATGATATCGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGACACACCGC    1368
Db      1324   GCTAGTATATCTGATATCAGATCAGATCCAGCGTGAATACGTTCCCGGGTCTTGACACACCGC    1383
Oy      1369   CCGTCAACACCATGAGAGTGGGTTTACCAAGAGAGGTAAGTCTA-ACCGTAAAGAGGAGC    1427
Db      1384   CCGTCAACACAGAGATTGTGAACCCGAAGTCGGTGGGGTACATCTACGGAGCCAG    1443
Oy      1428   CTTCGCAAGGTGATTCATGACTGGGGTG    1457
Db      1444   CCGCCGAAGGTGGAGCAGATGATTGGGGTG    1473

RESULT 36
US-11-055-637-79
; Sequence 79, Application US/11055637
; Publication No. US20050260619A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: DUBOIS, Jason
; APPLICANT: EDGE, Tom
; APPLICANT: MASSON, Luc
; APPLICANT: TREVORS, Jack T.
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
; FILE REFERENCE: 2139-33US
; CURRENT APPLICATION NUMBER: US/11/055,637
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,288
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 1513
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Probe for DNA array
US-11-055-637-79

```

Query Match	58.3%	Score 848.8;	DB 17;	Length 1513;
Best Local Similarity	77.3%	Pred. No. 5.2e-255;		
Matches 1136;	Conservative	0;	Mismatches 317;	Indels 17; Gaps 8

OY	5	AAGCGTGGCGGCAATGCTTTACATCCAGTCCGACGCGGAGCGTATCGTCACTCG	84
Db	1	AAGCGTGGCGGCGTCTTATATCATGCAGAGTCAGGAATGATTAAGACTTGCTCTTA	60
OY	65	TG--GCCAGTGGCGGACCGGTGAGTAAATGCATCGG--AACGTATCCGAGAAAGGGGGGTAA	121
Db	61	TGAAGTTAGCGGCGGAGCGGGTGAAGTAAACAGTGGGTAACTTACCCTAATAACTGGAGTTAA	120
OY	122	CGCATCCGAAAGATGCTGTAATACCGCAT--ATACTTAAGAGGAAAGACGGGGATCGAAA	180
Db	121	CTCCGGGAAACCGGGGCTAATATCCGGATATATTTTGAAGTGCATGATCTCGAAATTTGAAA	180
OY	181	GAC-----CTTGGCGCTTTTGGAGCGGCGCATGTCATGTAATGACTAGTTGGTGGGGTAA	232
Db	181	GGCGGCTTCGGCTGTCTCATTTATGGATGGAGCCCGGCTGCATTAAGTTGTTGGAGGTAA	240
OY	233	AGGCTTACCAAGCGGACGATCATAGTATGGTCTTGAGAGGACAGCACAGCCACACTCGGGACT	292
Db	241	CGGCTTACCAAGCGGACGATGCGTATGCCAGCTTGAGAGGGTGTATCGGCCACACTCGGGACT	300
OY	293	GAGACACGGGCCCAACCTCTTACCGGAGGACGACAGTGGGGAATTTTGGACATGGGCGCAA	352
Db	301	GAGACACGGGCCCAACCTCTTACCGGAGGACGACAGTGGGGAATCTTCCGCAATGGACGAAA	360
OY	353	GCCGATCCAGCAATGCGCGCTGAGTGAAGAAAGGCTTTCGGATTGTAAAGCTTTTCAGT	412
Db	351	GTCGACGGAGCAACGCCGCGTGAATGAAAGGCTTTCGGGTGTAAACTCTGTGTT	420
OY	413	CGAGAGAAAAGTTACCGTA--AATTAATCGTACTCATGACCGGTATGACAGAGAAAGCA	471
Db	421	AGGGAAAGAACAGTGTAGTTGAATAAGCTGGACCTTGAACGGTATCTTAACCAAGAAAGCC	480
OY	472	CCGGTAACTACGTGCGGACGAGCGCGGTAAATACGTAGGGGTGCAAGCGTTAATCGAATT	531
Db	481	ACGGTAACTACGTGCGGACGAGCGCGGTAAATACGTAGGTGCGAAGCGTTATCCGAATT	540
OY	532	ACTGGGCTTAAAGGATGCGCAGGCGGCTTTGTAAATGATGTAATGCCGAGTGAATCCCGGGCTTAA	591
Db	541	ATTGGGCTTAAAGCGCGCGGAGGTGTTCTTAAATGTCGATGTAAAGCCACACGGCTCAA	600
OY	592	CCTGGGAAATGGCTTTGAAACTACAGGCTTACAGTGTGGCAGAGGAGGTGGAAATTCAT	651
Db	601	CCGTGGAGGTTCATTTGAAACCTGGAGAGCTTGAAGTGCAGAAAGAGAAAGTGAATTCAT	660
OY	652	GTCGTAGCACTGAAAGCGTACAGATATGGAAGAAACATGATGGCGAAAGGCAAGCTCTCGG	711
Db	661	GTCGTAGCGGTGAAATGCGTACAGATATGGAAGAAACCAAGTGGCGAAGGCCACTTTCGG	720
OY	712	GTTAAACATGACGCTCATGCAAGAAAGCGTGGGAGCAACAGAGTTAGATACCTTGSTA	771
Db	721	TCTGTAACTGACACTGAGGCGGAAAGGCTGGGGAGCAAAACAGATTAAGATACCTTGSTA	780
OY	772	GTCACAGCCCTAAACGATGTCAACTAATTTGTTGGGCTTTATTAAGC--TTGTTAAAGAA	829
Db	781	GTCACAGCCGTAAACGATGATGTCTAATATGTTAAGGGTTTCCGCTTTATGTCGTAAAG	840
OY	830	CTAAGCGCGAAGTTGACCGGCTGGGGAGTACGAGTCGCAAGATTAAACTCAAGGAATT	889
Db	841	TTAAGCATTTAAGCACTCCGCTGGGGAGTACGCGCCGCAAGGCTGAAACTCAAGGAATT	900
OY	890	GACGGGAGCCCGCACAGCGGTGATTAATGTGATTTAATTCGATGCAACCGGAAAAACT	949
Db	901	GACGGGGGCGGCACAGCGGTGAGCAATGTGTTAATTCGAAGCAACCGGAAGAACT	960
OY	950	TACTTACCTTGCATGTAGCAATTTTCTAAGAGTAA--TTAAGTCTTGGGAAACGCTA	1008

Db	961	TACCAAGCTTGCACATCCTCTGAAAACTTGAAGATAAGAGCTTCTCTTCGGAGCACAG	1020
Qy	1009	ACACAGGTGCTGCATGAGCTGTCTGTCAAGCTCGTGTGTGAGATGTTGGGTAAATGCCGCA	1068
Db	1021	TGACAGGTGTGTGACATGTTGTGTGTCAAGCTCGTGTGTGAGATGTTGGGTAAATGCCGCA	1080
Qy	1069	ACGAGCGCAACCTTGTGATTAATTTGCAATCATTTTGGTGGCACCTTTAATGAGACTGCC	1128
Db	1081	ACGAGCGCAACCTTGTGATTAATTTGCAATCATTTTGGTGGCACCTTTAATGAGACTGCC	1140
Qy	1129	GGTGACAAACCGGAGGAAGTGGGGATGACGTCAGTCTCATGAGCCCTTAATGGGTAGGG	1188
Db	1141	GGTGACAAACCGGAGGAAGTGGGGATGACGTCAGTCTCATGAGCCCTTAATGGGTAGGG	1200
Qy	1189	CTTCACACGTATAACATATGCGCGCTTACAGAGGGTTCACCAACCCGCGAGGGGAGCTAATC	1248
Db	1201	CTACACACGTGTCAATATGACACGATACAAAGACTGTGCAAGAACCGGAGGTGGAGTAAATC	1260
Qy	1249	TCAGAAACCGCGTGTGAGTCCGGATTCGGAATCGAGTCTGCACTCGATCTCGTGAAGTGGAAATC	1308
Db	1261	TCATAAAAACCGTCTCAAGTTCGGAAATGTGAAGGTGTGCACTCCGCTTAATGAAGACTGGAAATC	1320
Qy	1309	GCTAGTATTCGCGGATTCAGCANTGTGCGGGGTGAATCGTTCGCCGGGTCTTGAACACACCGC	1368
Db	1321	GCTAGTATTCGCGGATTCAGCANTGCCGCGGTGAATCGTTCGCCGGGTCTTGAACACACCGC	1380
Qy	1369	CCGTACACCATGGGAGTGGGTTTCACCAAGACAGTATGTCATAAC-GTAAAGAGGGCG	1427
Db	1381	CCGTACACCAACGAGAGTGTGTAACACCCGAAGTGGGTGGGTAAACCTTTATGAGGACAG	1440
Qy	1428	CTTGCGACAGGTGAATTCATGACATCGGGGTG	1457
Db	1441	CCGCCTTAATGGTGGACAGATATATGGGGGTG	1470

RESULT 37  
US-11-055-637-80

```

: Sequence 80 Application US/11055637
: Publication No. US20050260619A1
:
: GENERAL INFORMATION:
: APPLICANT: BROUSSEAU, Roland
: APPLICANT: DUBOIS, Jason
: APPLICANT: EDGE, Tom
: APPLICANT: MASSON, Luc
: APPLICANT: TREVORS, Jack T.
: TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
: TITLE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES
: FILE REFERENCE: 2139-33US
: CURRENT APPLICATION NUMBER: US/11/055,637
: CURRENT FILING DATE: 2005-02-11
: PRIOR APPLICATION NUMBER: US 60/543,288
: PRIOR FILING DATE: 2004-02-11
: NUMBER OF SEQ ID NOS: 114
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 80
: LENGTH: 1505
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Probe for DNA array
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(1505)
: OTHER INFORMATION: n = A,T,C or G
: US-11-055-637-80

```

Query Match	58.1%	Score 845.8	DB 17	Length 1505
Best Local Similarity	75.8%	Pred. NO. 4.5e-254		
Matches 114; Conservative	0	Mismatches 340	Indels 16	Gaps 5

4 GAACGCTGGCGGCATGCTTACACATGCAGTGCAGCGCAGCACGGATGCTTGCACTCG 63

D	b	4	GAAGCTGGCGCGCTCTAATACATGCAATCCAGCGGACCAAGGAACTTGCTCCCA	63
O	y	64	GTGGGAGTGGCGGACGGGTGAGTAATGCAATCGG-AACTATCCAGAAAGAGGGGGTAACT	122
D	b	64	GAGGTTCACGGCGGACGGGTGAGTAACAGTGCNCAACTCCCTGTGACTGGGATTAAC	123
O	y	123	GCATCGAAGAGTGTCTTAATACCGCAT-----ATACTTAAGAGGAAACGAGGG	173
D	b	124	ATCGAAGAAATCGGTCTAATACCGGATATCAAGGAATCAATGTTCTTTGTAAAAAG	183
O	y	174	ATCGAAGAACCTTGGCGTTTGTGAGCGCGCATGTCTGATTAGCTAGTTGGTGGGTAAA	233
D	b	184	ATGGCTCCGGCTATACATCAGGAGGATGGCCCCCGCGCATTTAGTACTGTTGGTAAGTAT	243
O	y	234	GGCCTTACCAAGCGACGATCAGTAGTGGTCTGAGAGGACGACCACTGGGACTG	293
D	b	244	GGCTTACCAAGCGACGATGCGTAGCGCACTGAGAGGGTATGGGCGACACTGGGACTG	303
O	y	294	AGACACGGCCCCAGACTCTTACGGGAGCGACAGATGGGGAAATTTTGGACATAGGCGCAG	353
D	b	304	AGACACGGCCCCAGACTCTTACGGGAGCGACAGATGGGAATCTTCCGCAATGACGAAG	363
O	y	354	CCTGATCCAGAAATCCCGCGTAGTGAAGAAAGCCCTCGGGTTGAAAGCTCTTCAAGTC	413
D	b	364	TCTGACGAGGACCGCGCGTAGTGAAGAAAGTTTTCGAAATCGTAAGCTCTGTGTTTA	423
O	y	414	GAGAAAGAAAGTTTACGGT-AAATATATCGTACTCATATGACGGTATCGACAGAAAGCAC	472
D	b	424	GGAAAGAAACAAGTGCAGTGCATATAGTCCGCACTTGAACGTTACTTAACAGAAAGCA	483
O	y	473	CGGCTAATACGTGCAGACAGCCCGGTATATAGTAGGGTGCAGACGTTAATCGGAATTA	532
D	b	484	CGGCTAATACGTGCAGACAGCCCGGTATATAGTAGGGTGCAGACGTTGTCGGAATTA	543
O	y	533	CTGGGCGTAAAGGGGCGCAGCGCGCTTGTAACTCAGATGTGAATCCCGGGCTTAAC	592
D	b	544	TTGGGCGTAAAGCGCGCAGCGCGCTTTTAACTCAGATGTGAATTCGGGCTCAAC	603
O	y	593	CTGGGAATTTGGTTTGAACAATCAAGGCTAGAGTGTGGCAGAGGAGGTGAAATTCATG	652
D	b	604	CCGACAGGGGTATTTGAAACCTGGGAGACTTGAATCAAGAAAGAGATGGAAATTCACG	663
O	y	653	TGTAGCATGAATATGCGTAGATATGGAAGAACATGATGGCGAAGGCAAGCTTCCTGGG	712
D	b	664	TGTAGCGGTGAATATGCGTAGATATGTTGAGGAAACACAGTGGCCAAAGCGCACTCTGGT	723
O	y	713	TTAACATCGAAGCTCATGACGAAGAAAGGTGGGGAGCAACAGATTGATATCCTGGTAG	772
D	b	724	CTGTACTGACGTGAGGCGGGAAGCGTGGGGAGCAACAGATTGATATCCTGGTAG	783
O	y	773	TCCAGCGCCCTTAAAGATGTCAACTAGTTGTGTGGGCTTATATAGG--CTTGGTAACGAAC	830
D	b	784	TCCAGCGCGTAAAGATGATGATGCTTAGGTGTTTAGGGGTTTCGATCCCTTAGTGCGAAGT	843
O	y	831	TAAACCGTGAAGTTGACCGCGCTGGGGAGTACGTTGCAAGATTTAAAATCAAGAAATTG	890
D	b	844	TAAACATTTAAGCACTCGGCTGGGAGTAGCGCCGCAAGGCTGAATCTCAAGAAATTG	903
O	y	891	ACGGGGAACCGGCAACAGCGGTGATATATGTTGAATTTATTCGATGCAACCGCAAAAACCTT	950
D	b	904	ACGGGGGCGCCCAACAGAGTAGAGCATATGGTTTAAATTCGAAACCAACGCAAAACCTT	963
O	y	951	ACCTAACCTTGACATGTAGCAATTTTCTAGAGATAG--TTATGCTTTGGGAAACGCT	1007
D	b	964	ACCAAGTCTTACATCTTTTACCACTCTAGAGATAGAGCTTTCCTCTCGGGGGACAAA	1023
O	y	1008	AACAAGGTGTCGATGGCTGTCTGACCTGTGTCGTGAGATGTTGGGTTAAGTCCGCG	1067
D	b	1024	GTGACAGGTGTGATGGTGTGCTGACAGCTGTGTGTAAGATGTTGGGTTAATCTCCGC	1083
O	y	1068	AACGAGCGCAACCTTTGTCAATTAATTCGCATCATTTTGGTGGGCACTTAAATGAGACTGC	1127
D	b	1084	AACAGAGCAACCTTATATCTTAACTTGCAGACATTTAGTTGGGCACTCTAAGGTGACTGC	1143

Db 1084 AACGAGCCGACCCCTTGATCTTAGTTGCCAGCATTTAGTTGGGCACTTAAGGTGACCTGC 1143

Qy	1128	CGGTGACAAACCCGAGGAAAGTGGGGATGACGTCAAGTCCGATAGGCCCTTATGSGTAAAG	1187
Db	1144	CGGTGACAAACCCGAGGAAAGTGGGGATGACGTCAAAATCAATCATGCCCCCTTATACACTGG	1203
Qy	1188	GCTTCACACGTATATCAATATGGCGGTCACAGAGGGTTGCAACCCGCGAGGGGAGCTATAT	1247
Db	1204	GCTACACACGTGCTACATATGATGTGTACAAAGGGCAGCGAATCCGCGAGGTGAGGCCAAT	1263
Qy	1248	CTCAGAAACCGCGTCTGATGCCGATCCGATCTGCAACTCGTGAAGTCGGAT	1307
Db	1264	CCCATAAAGCCATTCTCAGTTTCGATTGTAGGCTGCAACTGGCTCATATGAAGCCGGAT	1323
Qy	1308	CGTATATATCGCGGATCAGCATGTCCGCGTGAATAGTTCCCGGGTCTTGTACACACC	1367
Db	1324	TGCTATGTAATCGCGGATCAGCATGTCGCGTGAATACGTTCCCGGGCTTGTACACACC	1383
Qy	1368	CCCGTCACACCATGGGATGGGTTTACACAGAGAGGTAGCTTAAACCGTAAAGAGGGCG	1427
Db	1384	CCCGTCACACACGAGAGTGTGTAAACACCGAAGTCGGTGGGATTAACCTTTTGGAGCCAG	1443
Qy	1428	CTTGCCACGGTGAATCATATGCTGGGGG	1457
Db	1444	CGGCTTAAGTGGACAGATGATTTGGGGTG	1473

```

RESULT 38
US-11-273-617-6
; Sequence 6, Application US/11273617
; Publication No. US20060067924A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Margie D.
; APPLICANT: Hofacre, Barry
; APPLICANT: Harmon, Barry
; TITLE OF INVENTION: Probiotic Bacteria and Methods
; FILE REFERENCE: 18-03
; CURRENT APPLICATION NUMBER: US/11/273,617
; CURRENT FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/470,807
; PRIOR FILING DATE: 2003-05-14
; PRIOR APPLICATION NUMBER: WO us2004/015378
; PRIOR FILING DATE: 2005-04-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Enterococcus cecorum
US-11-273-617-6

```

Query Match	57.7%	Score 841.2;	DB 18;	Length 1509;
Best Local Similarity	77.8%	Pred. No. 1.2e-252;		
Matches 1090; Conservative	0;	Mismatches 299;	Indels 12;	Gaps 6.

Qy	69	GAGTGGCGGACGGGTAGTAATGCAATCGG-AA	CGTATCCAGAAAGAGGGGGTAACGCATC	127
Db	81	GAGTGGCGAAGGGGTAGTAACACGTGGGTAACT	CGCCCATCAGCGGGGGATTAACA	140
Qy	128	GAAGATGTGCTTAATACCGCAT-AT	ACTCTAAAGAGAAAGCAGGGATCGAAAGCCTT	186
Db	141	GAAACAGGTGTCTAATACCGATTAATTC	CAATTTACCGCATGTATGATGATGAAGGCGCT	200
Qy	187	G-----GCCTTTGGAGCGCGCGAAGTCTG	ATTAAGTTGGTGGGGTTAAAGGCTAC	240
Db	201	TTTTCGTCACTAGTAGTAGAATGCCGCGTGC	ATTAAGTTGGTGGGGTTAAAGGCTAC	260
Qy	241	CAGGCGACGATCAGTAGTTGTCTGAGAGAC	CGACCACACTTGGGACTGACACACG	300
Db	261	CAGGCTCGCGATTCATATACCGACTTAGAG	GGGTAGTCGGCACACTGGGACTGAGACAG	320
Qy	301	GCCCAAGCTCTTACGGGAGGCACAGTGGGGA	TTTTGGACAATTGGGCGCAAGCCTGATC	360
Db	321	GCCCAAGCTCTTACGGGAGGCACAGTGGGGA	TTTTGGACAATTGGGCGCAAGCCTGATC	380

QY	361	CAGCAATGCCGCGTGAATGAAGAAGCCTTCGGGTGTAATAGCTTTTCAGTGGAGAGA	420
Db	381	GAGCAACGCCGCGTGAATGAAGAAGTTTTCGATGCTAAACCTCTGTGTAGAGAGA	440
QY	421	A-AAGGTTAAGGTAAATATATCGTACCTCATGACGATTCGACAGAGAACACCGGCTAA	479
Db	441	ACAAGATGAGAGTGAAGAAATTTCATCCCTTGACGATCTTAAACGAAAGCCAGGCTAA	500
QY	480	CTACGTCCAGACAGCCCGGTAATACGTAGGGTGCAAGCCTTAATCGAAATTAAC	539
Db	501	CTACGTCCAGACAGCCCGGTAATACGTAGGGTGCAAGCCTTGACGATTTTAATGCGGC	560
QY	540	TAAAGGTGCGCAGCGCGCTTTGTAAGTCAGATGTGAATATCCCGGGCTTAACCTGGAA	599
Db	561	TAAAGCGAGCCAGCGCGCTTTTAAGTCAGATGTGAATATCCCGGGCTTAACCGGGAG	620
QY	600	TTGGCTTTGAATCTAACAAAGCTAAGTGTGCAGAGGAGAGGTGAATTCATGTTGTACA	659
Db	621	GGTCAATTGAAACCTGGGAGACTTAAGTCAGAAAGAAAGGAAATTCATGTTGTAGCG	680
QY	660	GTAAGATCCGAGAGATATGGAGAAATCGATGCGCAAGCAGCTTCGGATTAAAC	719
Db	681	GTAAGATCCGAGATATATGAGAGAACACAGTGGCGAAGCGGCTTCTGTGCTGTAAAC	740
QY	720	TGACGCTCATSCACGAAACGCTGGGAGCAAAACGATTAGATATCCTGTTATGTCACGC	779
Db	741	TGACGCTCATGCTCGAAAGCGTGGGAGCAAAACGATTAGATATCCTGTTATGTCACGC	800
QY	780	CCTAAACGATGTCACTATGTTGTTGGGCTTAATAGGCTTG - GTAACGAAGCTAACGCG	837
Db	801	CGTAAACGATAGTCTTAAGTTGGAGGGTTTCGGCTTCAGTGTCTCAGCAAAACGCA	860
QY	838	TGAATGTACCGCCTGGGAGTACGATCGATCGCAAGATTAAACTCAAGAAATTGACGGGA	897
Db	861	TTAAGCACTCCGCTGGGAGTACGACCGCAAGGTTGAAATCAAGGAAATTGACGGGA	920
QY	898	CCCGCACAAGCGGTGATTAATGTGGATTAAATTCGATGCACGCGAAAAACTTACCTAAC	957
Db	921	CCCGCACAAGCGGTGATTAATGTGGATTAAATTCGATGCACGCGAAAGAACTTACCTAAC	980
QY	958	CTTGAATGTAGCAATTTTCTTAGAGATTA - GATTAGTCTTCGGGAACGCTAACACAGGT	1016
Db	981	CTTGAATGTAGCAATTTTCTTAGAGATTAATTTCCCTTCGGGGAACAAAGTACAGGT	1044
QY	1017	GCTGATGTGCTGTGATGCTGATGCTGATGATGTTGGTTAAGTCCCGCAACGAGCGC	1076
Db	1041	GCTGATGTGCTGTGATGCTGATGCTGATGATGTTGGTTAAGTCCCGCAACGAGCGC	1100
QY	1077	AACCTTGTCTTAATTTGCCATCAATTTGTTGGGCACTTAAATGACCTGCGCGTGACAA	1138
Db	1101	AACCTTGTCTTAATTTGCCATCAATTTGTTGGGCACTTAAATGACCTGCGCGTGACAA	1160
QY	1137	ACCGAGGAAGGTGGGATGACGTCAAGTCTCATGCGCTTAATGAGGATTCACAC	1196
Db	1161	TCGAGGGAAGGTGGGATGACGTCAATTCATCAAGCCCTTATACCTGGGCTTACACAC	1222
QY	1197	GTAATATCAATGCGCGGTACAGAGGTTGCCAAACCGGAGAGGGGAGCTAATCTCAAGAA	1256
Db	1221	GTAATATCAATGAGAGTACCAAGATTCGAAAGCCGCGAGGCTTAAGCCAAATCTTTAAAG	1288
QY	1257	CGGCTGTATGCTCGGATCGGAGTCTGCACCACTCGATCGTGAAAGTCGAAATGCTAGTAA	1316
Db	1281	CTCTTCTCAGTTCCGATTTGTAGGCTGCACCACTCGCTTACATGAAGCCGGAATGCTAGTAA	1344
QY	1317	TGCGGATATCAAGATTCGCGATGAATACGTTCCCGGCTTTGTACACACCGCGCTCACAC	1376
Db	1341	TGCGGATATCAAGACCGCGGATGAATACGTTCCCGGCTTTGTACACACCGCGCTCACAC	1400
QY	1377	CCATGGAGATGGGTTTACACAGAACCAAGTATGCTAAACGTAAGGAGGCGCTTGCCACG	1436
Db	1401	CCACGAGATTTGTAAACACCAACACCGGATCGGTAAACCGCAGGAGCCAGCGCTTAAG	1466









Query Match 57.6%; Score 838.6; DB 18; Length 1587;  
 Best Local Similarity 76.7%; Pred. No. 8.2e-252;  
 Matches 1075; Conservative 0; Mismatches 314; Indels 12; Gaps 4;

```

69 GAGTGGCGGACGGGTGAGTATGATCGG-AACTATCCAGAAAGGGGGGTAAAGCATC 127
121 GAGTGGCGGACGGGTGAGTATGATCGG-AACTATCCAGAAAGGGGGGTAAAGCATC 180
128 GAAAGATGTGCTTAATCCGATATCTCTAAGAGAGAAAGCGGGGATCGAAAGC---- 183
181 GAAAGATGTGCTTAATCCGATATCTCTAAGAGATGATGATCTTAAGTAAAGATGAT 240
184 ----CTTGGGCTTTTGGAGCGGCGGATGCTGATTAAGTATGATGATGATGATGAT 239
241 TCTGCTATGCTTTTGAATGAGACCGCGGCTTATTAAGTATGATGATGATGATGAT 300
240 CCAAGGCGACGATCAGTATGCTGCTGAGAGACGACGACCACTGGGACTGAGACAC 299
301 CCAAGGCTATGATAGTATGAGCGGACGATGAGAGGTTGATGGGCACTGGGACTGAG 360
300 GGCCGACATCTCTACGGAGAGGACGATGAGGGAATTTTGAACATGGGCGCAAGCTGAT 359
361 GGTCCAAACTCTCTACGGAGAGGACGATGAGGGAATTTTCACAATGAGACGCAAGCTGAT 420
360 CCAGCAATGCGGCGGATGAGAAAGGCGCTTGGGTTGTAAGCTCTTCACTGATGAGAG 419
421 GGAGCAAGCGGCGGATGAGAAAGGCGCTTGGGTTGTAAGCTCTTCACTGATGAGAG 480
420 AAAAGGTTACGGTAAATATCTGATCTGATGATGATGATGATGATGATGATGATGAT 479
481 AACAGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
480 CTACGTGCGACGACCGCGGTAAATCTGATGATGATGATGATGATGATGATGATGAT 539
541 CTACGTGCGACGACCGCGGTAAATCTGATGATGATGATGATGATGATGATGATGAT 600
540 TAAAGGTTACGGTAAATATCTGATCTGATGATGATGATGATGATGATGATGATGAT 599
601 TAAAGGTTACGGTAAATATCTGATCTGATGATGATGATGATGATGATGATGATGAT 660
600 TTGCGTTTGAATCTCAAGGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTAGCA 659
661 GTGATTTGGAATCTGGAAGACTGATGATGATGATGATGATGATGATGATGATGAT 720
660 GTGAAATCTGTAAGATATGGAAGAACTCGATGCGGAAGCGCTCTGGGTTAAAC 719
721 GTGAAATCTGTAAGATATGGAAGAACTCGATGCGGAAGCGCTCTGGGTTAAAC 780
720 TGAGGCTGATGACGAAAGCGTGGGAGCAACAGATTAAGTATGATGATGATGATGAT 779
781 TGAGGCTGATGACGAAAGCGTGGGAGCAACAGATTAAGTATGATGATGATGATGAT 840
780 CCTAAACGATGCACTAGTTGTTGGGCTTATTAAGGCTTG--GTAAAGAACTAAACGCG 837
841 CGTAAACGATGCACTAGTTGTTGGGCTTATTAAGGCTTG--GTAAAGAACTAAACGCG 900
838 TGAAGTTAAGCGGCTGGGAGATGCGTGGCAAGATTAAGTAAAGTAAAGTAAAGTAAAG 897
901 ATAAGCATTTCCGCTGGGAGATGACGCAAGGTGTAAGTAAAGTAAAGTAAAGTAAAG 960
898 CCCGCAAGAGGGGATGATGATGATTAATTCATGATGATGATGATGATGATGATGAT 957
961 CCCGCAAGAGGGGATGATGATGATTAATTCATGATGATGATGATGATGATGATGAT 1020
958 CTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016
1021 CTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1017 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
1081 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140

```

```

1077 AACCTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136
1141 AACCTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1137 ACCGAGAGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196
1201 ACCGAGAGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1197 GTATTAAGTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256
1261 GTGCTAAGTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1257 GCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1316
1321 CCGTTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1317 TCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1376
1381 TCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1377 CCATGAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1436
1441 CCATGAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1437 GTGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
1501 GTGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1521

```

RESULT 42  
 US-11-055-637-71  
 ; Sequence 71, Application US/11055637  
 ; Publication No. US20050260619A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BROUSSEAU, Roland  
 ; APPLICANT: DUBOIS, Jason  
 ; APPLICANT: EDGE, Tom  
 ; APPLICANT: MASSON, Luc  
 ; APPLICANT: TREVORS, Jack T.  
 ; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND  
 ; TITLE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES  
 ; FILE REFERENCE: 2139-33US  
 ; CURRENT APPLICATION NUMBER: US/11/055,637  
 ; CURRENT FILING DATE: 2005-02-11  
 ; PRIOR APPLICATION NUMBER: US 60/543,288  
 ; PRIOR FILING DATE: 2004-02-11  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 71  
 ; LENGTH: 1486  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Probe for DNA array  
 US-11-055-637-71

Query Match 57.5%; Score 837.2; DB 17; Length 1486;  
 Best Local Similarity 75.9%; Pred. No. 2.2e-251;  
 Matches 1117; Conservative 0; Mismatches 318; Indels 17; Gaps 6;

```

3 TGAAGCTGCGGCGGATCTTAAACATGCAAGTCAAGC-GCAGACGATGCTTGATC 61
3 TGAAGCTGCGGCGGATCTTAAACATGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 62
62 T-GGTGCGAGTGTGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 119
63 TATGACGTATGCGGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 122
120 AACGATCGAAAGATGCTTAATACCGCATATCTCTAA-----GGAAGAAAGCAG 170
123 AACTTGGGAAACCGAGCTTAATACCGCATATCTCTCTTCAATGAGAGATGATGA 182
171 GGATCGAAAGACTTTCGCTTTTGGAGCGCGCATGCTGATTAAGTATGATGATGATGAT 230

```

Db	183	AAGATGGTTTCGGCTATCACTTACAGATGGGCCCGCGGTGCATTAGCTAGTTGGTAGGCT	242
Oy	231	AAAGGCTTACCAAGGCGACGATCAGTAGTTGGTCTGAGAGACGACCACTGGGA	290
Db	243	AAACGGCTCACCAAGGCAACGATGGGTAGCCGACCTGAGAGAGGTGATCGGCCACACTGGGA	302
Oy	291	CTGAGACACGGCCGACACTCTTACGGGAGGACGACGTGGGGAATTTTGGACAAATGGCGC	350
Db	303	CTGAGACACGGCCGACACTCTTACGGGAGGACGACGTAGGGAATCTTCCCAATGGACGA	362
Oy	351	AAAGCTGATCCAGCAATGCCCGGTGAGTGAAGAAAGGCGTTGGGTGTGAAGCTCTTCA	410
Db	363	AAAGCTGACGGAGCAAGCGCCGTGAGTGAAGAAAGGCGTTTGGGTGTGAAGCTCTTGGT	422
Oy	411	GTCCAGAAAGAAAGTTTACGGTAAATATACGTGACTCATGACGGTATTCGACAGAAAGC	470
Db	423	TTTAGGAAAGAACAAAGTACAAAGATGATCTGCTTGTACTTGAACGTTACTTACCAAGAAAGC	482
Oy	471	ACCGGCTTAATACGTGCCAGCAGCCCGCGGTAAATACGTAGGGTGCAGCGCTTAATCGGAAT	530
Db	483	CACGGCTTAATACGTGCCAGCAGCCCGCGGTAAATACGTAGGGTGCAGCGCTTAATCGGAAT	542
Oy	531	TACTGGCGGTAAAGGATGGCGGACGGCGCTTTGTAAAGTCAAGTGGAAATCCCGGGCTTA	590
Db	543	TATTGGCGGTAAACCGCGCGGACGGCGCTTTGTAAAGTCAAGTGGAAATCCCGGGCTTA	602
Oy	591	ACCTGGGAAATTGCGTTTAAACTCAAGAGGCTAGAGTGGCGACAGAGGATGGAATTCGA	650
Db	603	ACCGTGAAGGGTCAATTGGAAATCTGGGGAACCTTGAGTCCAGAAAGAAAGGGAATTCGA	662
Oy	651	TGTGTAGCAGTGAATGCGTAGAGATATGAAAGCAATCGATGGCGAAGGCAAGCTCTCTG	710
Db	663	CGTGTAGCGGTGAAGATGCGTAGAGATATGAGAGCAACAGTGGCGAAGGCGGCTTTTGG	722
Oy	711	GCTTAACTGACGCTCATTCGACCAAGAAAGCTGGGGAGCAACAGAGTTAGTAACTCCGTG	770
Db	723	GTCTGTAACTGACGCTGAGGGCGCAAGACGGGGGAGCAACAGAGTTAGTAACTCCGTG	782
Oy	771	AGTCCAGCCCTTAAACGATGTCACTAGTTGTGGGCGCTTATTTAGGC--TTGTAAACGA	828
Db	783	AGTCCAGCCCTTAAACGATGTCACTAGTTGTAGAGGTTTCCGCCCTTATAGTCTGCA	842
Oy	829	GCTAAACGGGTGAAGTTGAACCGCTGGGGAGTACGGTGCAGAGATTAATACTCAAGGAAT	888
Db	843	GCTAAACGGATTAAGACATCCGCTGGGAGTACGGTGCAGAGATTAATACTCAAGGAAT	902
Oy	889	TGACGGGGACCCGCAACAGCGGTGATTAATGTGATTAATTCGATGCAAGCCGAAAGAAC	948
Db	903	TGACGGGGGGCCGCAACAGCGGTGAGCATGTGTTTAATTCGAAGCAACCGGAAGAAC	962
Oy	949	TTACCTACCCCTTGAATGAGCAATTTCTTCAAGATAGA---TTAGTCTCGGGAACG	1005
Db	963	TTACCAAGGCTTGAATGAGCAATCTTGAACACTTAAGAGATAGAGCGTTCCCTTGGGGGACA	1022
Oy	1006	CTAACAACAGTGTCTCATGTGCTGTCAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCC	1065
Db	1023	GAGTGAACAGTGTGTGATGTTGTGTGTCAGCTCGTGTGAGATGTTGGGTTAAGTCCC	1082
Oy	1066	GCAACGAGGCGAACCCCTTGTCACTTAATTGGCAATCTTTGGTGGGCACTTTAATGAGACT	1125
Db	1083	GCAACGAGGCGAACCCCTTGTCACTTAATTGGCCAGCACTTAAGTGGGCACTTAAAGTGA	1142
Oy	1126	GCCGGTGAACAACCGGAGGAGAGTGGGGATGACGACAGTCTCAATGGCCCTTAATGGGTA	1185
Db	1143	GCCGGTGAACAACCGGAGGAGAGTGGGGATGACGACCAATCACTATGCGCCCTTATGACT	1202
Oy	1186	GCGCTTCAACAGTAAATACATGTCGCGGTACAGAGGTTGGCCAAACCGCGAGGGGAGCTA	1245
Db	1203	GGGCTTCAACAGTGTGCTAATGATGTGTAACAAGGGCTGCAAGACCGGAGGCTCAAGCA	1262
Oy	1246	ATCTCAAAAAGCGCGTCTGATGTCGGATTCGAAGTCTGCACTTCGACTCCGTGAAGTCGGA	1305

```

Db      1263 ATCCATATAAACCATTTCTCAGTTCGGATTGTATAGGCTGCAACATCGCTATACATGAAGCTGGA 1322
Qy      1306 ATCGTAGTAATATCGCGGATATCAGCATATGTCGGGTGAATACGTTCCGGGGTTTGTATACACAC 1365
Db      1323 ATCGTAGTAATATCGCGGATATCAGCATATGTCGGGTGAATACGTTCCGGGGTTTGTATACACAC 1382
Qy      1366 CGCCCGTCAACACATGGAAGTGGGTTTCACCAAGACAGTGTCTTAACCGTAAAGAGGG 1425
Db      1383 CGCCCGTCAACACAGAGAGTTTGTAAACACCCGAATCGGTGAGTAAACCGTAAGGACGT 1442
Qy      1426 CGCTTGCCACGGGTGAGATTATGATGCTGGGGTG 1457
Db      1443 AGCCGCTTAAGTGGGACAGATGATTGGGGTG 1474

RESULT 43
US-11-055-637-78
; Sequence 78, Application US/11055637
; Publication No. US20050260619A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: DUBOIS, Jason
; APPLICANT: EDGE, Tom
; APPLICANT: MASSON, Luc
; APPLICANT: TREVORS, Jack T.
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
; TITLE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES
; FILE REFERENCE: 2139-33US
; CURRENT APPLICATION NUMBER: US/11/055,637
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,268
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4. 0
; SEQ ID NO 78
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Probe for DNA array
US-11-055-637-78

Query Match      57.5%; Score 837.2; Db 17; Length 1507;
Best Local Similarity 76.0%; Pred. No. 2.2e-251;
Matches 1116; Conservative 0; Mismatches 333; Indels 19; Gaps 6;

Qy      5 AACGCGGCGGCGCATGCTTTTACATCATCAAGTCGAAACGGGCGACACGATGCTTGCATCTGG 64
Db      1 AACGCTGGGGGCGGCTGCTTAATACATCAATCGAATCGAGC-----GAATGTTGAGTTTAActCA 56
Qy      65 TGGCGAGTGGCGGACGGGTGAGTAATGATCATCG- AACGTATCCAGAAAGGGGGGTAAACG 123
Db      57 CAATTAGCGGCGGACGGGTGAGTAATCAACGCGGCAACCTCGCTTAAGACTGGGATTAAC 116
Qy      124 CATGAAAGATGTGCTAATATCCGCATATATCTTAAGAGAGAAAG-CAGGGATGAGAAAGA 182
Db      117 TCGGGAACCGGAGCTAATATCCGATATGTTCTTCTTCGATAGAGAAATGGAGAA 176
Qy      183 C-----CTTGCAGCTTTTGGAGCGGCGGATGCTGATTAGTATGTTGGTGGGTAAAG 234
Db      177 CGGTCTCGGCTGTCACTTAATAGATGGGCCCGCGCGCATTAAGTATGTTGGTGAATATG 236
Qy      235 GCCTTACCAAGGCGAGCATAGTATGTTGGTGAAGAGACGACCGACCACTTGGGACTGA 294
Db      237 GCTCAACCAAGGCAACATGCTGAGCCGACCTGAGAGGGGTGATTCGGCCACACTGGGACTGA 296
Qy      295 GACACGGCCCAAGCTCTTACGGAGGCGAGAGTGGGGAATTTTGAACAATGGGCGCAAGC 354
Db      297 GACACGGCCCAAGCTCTTACGGAGGCGAGAGTGGGGAATTTTCCGCAATGAGAGAAAGT 356
Qy      355 CTGATCATCAGCAATGCGCGCTGAGTGAAGAAAGGCTTGGGTGTAAAGCTTTTCAGTCG 414
Db      357 CTGACGAGAGAAACGCGCGGTGAAGAAAGGCTTGGGGTGTAAAGTTCGTGTGTAG 416

```



```

Db      ||||| 2157 GTGTAGCCGTAAATGCGCAGATATGAGGAAACACAGTGGCGAAGCCGACTTCTGG 2098
Qy      ||||| 712 GTTAACTGACCGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGGTA 771
Db      ||||| 2097 TCTGTAATCTGACGCTGATGTGCGAAGCGTGGGATCAAAAGATTAGATACCTGGTA 2038
Qy      ||||| 772 GTTCAAGCCCTTAAACGATGTCAATAGTTGTT--GGGCTTAAATTAGGCTTGGTAAAGAA 829
Db      ||||| 2037 GTTCAAGCCCTTAAACGATGTGCTAAGTGTAGGGGGTTTCGCGCTTAACTAGTGTGAG 1978
Qy      ||||| 830 CTAAACCGTGAAGTTGACCGCTGGGAGTACGCTGCGAAAGTTAAACTCAAGAAATT 889
Db      ||||| 1977 CTAAAGCATTAAGCACTCCGCTGGGAGTACGACCGAAAGGTTAAACTCAAGAAATT 1918
Qy      ||||| 890 GACGGGAGCCCGCAAGCGGTGATGATTAATGATTAATGATGCAACGCAAAACCT 949
Db      ||||| 1917 GACGGGAGCCCGCAAGCGGTGAGCATGTGTTTAATTCGAAGCAACGCAAAACCT 1858
Qy      ||||| 950 TACCTACCTTGAATGATGAGCAATTTTCTAGAGATAGATTAGTGC--TTGCGGAAACGC 1006
Db      ||||| 1857 TACCAATCTTGAATCTCTCTGACCCCTCTAGAGATAGATTTCCTTGGGGAGACG 1798
Qy      ||||| 1007 TAAACAAGTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1066
Db      ||||| 1797 AGTAACAGGTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1738
Qy      ||||| 1067 CAACAGCGCAACCTTGTCTAATTAATGATGATTAATGATGATGATGATGATGATGATG 1126
Db      ||||| 1737 CAACAGCGCAACCTTGAAGCTTGAAGTGCATTAATTAATTAATTAATTAATTAATTAAT 1678
Qy      ||||| 1127 CCGGTGACAAACCGGAGGAGTGGGATGAGTCAAGTCTCATGACCCTTAATGAGTAG 1186
Db      ||||| 1677 CCGGTGACAAACCGGAGGAGTGGGATGAGTCAAGTCTCATGACCCTTAATGAGTAG 1618
Qy      ||||| 1187 GCGTTCACAGTAAATCAATGCGCGCTGACAGAGGTTGCCAACCCCGAGGGGAGCTAA 1246
Db      ||||| 1617 GCGTTCACAGTAAATCAATGCGCGCTGACAGAGGTTGCCAACCCCGAGGGGAGCTAA 1558
Qy      ||||| 1247 TCTCAGAAAGCCGCTGCTGATGCTGCGATGCGAAGTCTGCACTCCGTAAGTGGAA 1306
Db      ||||| 1557 TCCCAATAAAGTTGTTCTCAGTTGCTGATGATGATGATGATGATGATGATGATGATG 1498
Qy      ||||| 1307 TCGCTAGTATCGCGGATGAGCATGTCGCGTGAATACGTTCCCGGCTTGTACACACC 1366
Db      ||||| 1497 TCGCTAGTATCGCGGATGAGCATGTCGCGTGAATACGTTCCCGGCTTGTACACACC 1438
Qy      ||||| 1367 GCCCGTCAACCAATGGAGTGGGTTTCAACCAAGAGAGTGTCTAACCGTAAAGAGGC 1426
Db      ||||| 1437 GCCCGTCAACCAATGGAGTGGGTTTCAACCAAGAGAGTGTCTAACCGTAAAGAGGC 1378
Qy      ||||| 1427 GCTTGCCACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1457
Db      ||||| 1377 GCCGTGAAGGTGGGACCAATGATGGGGTG 1347

```

```

RESULT 45
US-10-793-626-3905
; Sequence 3905, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3905
; LENGTH: 3308
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3905

```

```

Query Match      57.4%; Score 837; DB 10; Length 3308;
Best Local Similarity 75.5%; Pred. No. 3,1e-251;
Matches 1110; Conservative 0; Mismatches 345; Indels 16; Gaps 5;

```

```

Qy      ||||| 3 TGAACGCTGCGCGGATGCTTATACATGCAATGCAAGTGAACGCGACGACGATCTTGATCT 62
Db      ||||| 235 TGAACGCTGCGCGGATGCTTATACATGCAATGCAAGTGAACGCGACGACGATCTTGATCT 294
Qy      ||||| 63 GGTGCG--AGTGGCGACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 119
Db      ||||| 295 TCTGACGTTACCGCGGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 354
Qy      ||||| 120 AACGATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 171
Db      ||||| 355 AACCTGCGGAAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 414
Qy      ||||| 172 GATGAAAGACCTTGGCTTGGAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 231
Db      ||||| 415 AAGACGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
Qy      ||||| 232 AAGGCTTACCAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 291
Db      ||||| 475 ACGGCTTACCAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 534
Qy      ||||| 232 TGAACGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351
Db      ||||| 535 TGAACGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 594
Qy      ||||| 352 AGCTGATGACGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 411
Db      ||||| 595 AGCTGATGACGCAATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 654
Qy      ||||| 412 TGAAGAAAGGTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 471
Db      ||||| 655 TGAAGAAAGGTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 714
Qy      ||||| 472 CCGGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 531
Db      ||||| 715 ACGGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 774
Qy      ||||| 532 ACTGAGCGTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 591
Db      ||||| 775 ATTGGGCGTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 834
Qy      ||||| 592 CCTGGGATTTGCTTGAATCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 651
Db      ||||| 835 CCGTGAAGGCTTATGAAATGAAACTTGAAGTCAAGAAAGGATGATGATGATGATGATGATGATGATG 894
Qy      ||||| 652 GTGTAGCATGAAATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 711
Db      ||||| 895 GTGTAGCATGAAATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 954
Qy      ||||| 712 GTTAACTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 771
Db      ||||| 955 TCTGTAATCTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
Qy      ||||| 772 GTTCAAGCCCTTAAACGATGTCAATAGTTGTT--GGGCTTAAATTAGGCTTGGTAAAGAA 829
Db      ||||| 1015 GTTCAAGCCCTTAAACGATGTCAATAGTTAGGGGGTTTCGCGCTTAAAGTGTGACAG 1074
Qy      ||||| 830 CTAAACCGTGAAGTTGACCGCTGGGAGTACGCTGCGAAAGTAACTCAAGAAAGATT 889
Db      ||||| 1075 CTAAACCGTGAAGTTGACCGCTGGGAGTACGCTGCGAAAGTAACTCAAGAAAGATT 1134
Qy      ||||| 890 GACGGGAGCCCGCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 949
Db      ||||| 1135 GACGGGAGCCCGCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1194

```

Oy		950	TACCTACCCTTGCACATGTAGCGAATTTCCTGAGATGATTAGTGC---	TTCCGGMAACC	1006
Db		1195	TACCAATCTTGACAATCCTCTGACCCCTCTGAGAATAAGTTTTCCCCTCGGGGACAG		1254
Oy		1007	TAACACAGGTCCTGCATGGCTGTGCGTCACGTCGTCGTGCGTAGATGTTGGGTTAAAGTCCCG		1066
Db		1255	AGTAAACAGGTGTGCATGGTGTGTGCTCACGTCGTGTGCTGAGATGTTGGGTTAAAGTCCCG		1314
Oy		1067	C AACGAGCGCAACCTTGTCTATTAATTGCCATCATTTGGTGGGCACTTTAATGAGACTG		1126
Db		1315	CAACGAGCGCAACCTTGAAGCTTAGTTGCCATCATTAAGTTGGGCACTCTAAGTTGACTG		1374
Oy		1127	CCGGTGAACAAACCGGAGGAAGTGGGGATGAGACGTCAAGTCCCTCAATGGCCCTTATGGGTAG		1186
Db		1375	CCGGTGAACAAACCGGAGGAAGTGGGGATGAGACGTCAATCAATCAATGCCCCCTTAATGATTTG		1434
Oy		1187	G GCTTCAACAGCTAAATACAAATGGCGCGTACAGAGGGTTGCCAACCCCGCAGGGGGAGGTAA		1246
Db		1435	G GCTACACACGTCGTCAATAGGA CAATCAAAGGGCAGCAACCGGAGGTTCAAGCAAA		1494
Oy		1247	TCTCAGAAAAGCGCGTCGTAGTCCGGATGGGAGTCTGCAACTCGACTCCGTGAATCCGGA		1306
Db		1495	TCCCATTAAGTTGTTCTCAGTTCCGATTGTAGTCTGCAACTCGACTCATATGAAGACTGAAA		1554
Oy		1307	TCGCTAGTAATCGCGGATCACAGTCGCGGTGATAAGTTCCCGGGCTTGTACACACC		1366
Db		1555	TCGCTAGTAATCGTAGATCACAGTCGTACGGGTGAATACGTTCCGGGGCTTGTACACACC		1614
Oy		1367	G CCGCTCACACCATGGAGAGTGGGTTTCCACCAAGAACAGTAGTCTAACCGTAAAGAGGCG		1426
Db		1615	G CCGCTCACACCAAGAGAGTTGTGAACACCCGGAAGCGGTGAGTAACCATTTGGAGCTA		1674
Oy		1427	G CTTGCCACGSGTAGATTCATGACTGGGGTG		1457
Db		1675	G CCGTCAAGSTGGGACAAATGATTTGGGGTG		1705
RESULT 46					
US-10-793-626-4187/c					
; Sequence 4187, Application US/10793626					
; Publication No. US20050255478A1					
GENERAL INFORMATION:					
APPLICANT: KIMBERLY, WILLIAM JOHN					
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS					
FILE REFERENCE: PUS480US					
CURRENT APPLICATION NUMBER: US/10793.626					
PRIORITY FILING DATE: 2004-03-04					
PRIOR APPLICATION NUMBER: 60/164,258					
PRIORITY FILING DATE: 1999-11-09					
NUMBER OF SEQ ID NOS: 4472					
SOFTWARE: PatentIn Ver. 2.1					
SEQ ID NO 4187					
LENGTH: 3657					
TYPE: DNA					
ORGANISM: Artificial Sequence					
FEATURE:					
OTHER INFORMATION: Description of Artificial Sequence: synthetic					
US-10-793-626-4187					

Query Match	57.4%;	Score 837;	DB 10;	Length 3657;
Best Local Similarity	75.5%;	Pred. No. 3.2e-251;		
Matches 1110;	Conservative 0;	Mismatches 345;	Indels 16;	Gaps 5;
QY	3	TGAACGCTGGCGCATGCTTTACACATGCAAGTGGAAACGACGACGAGTCTTGCACTCT	62	
Db	1625	TGAACGCTGGCGCGGTGCTCTAATCATGCAAGTGCACGAAACAGACGAGGACTTGCTCC	1566	
QY	63	GGTGGCC--AGTGGCGGACGGGTGATGATCATCGG--AACGTATCCAGAAAGAGGGGGCT	1119	
Db	1565	TCTGACCTTACGGCGCGACGGGTAGTAAACAGTGGATTAACCTTACCTTAAAGACTGGGAT	1506	

QY	120	TAAGCATCGAAGAAGATGCGTAAATCCCGAT- - - - -TACTTAAGAGAGAAAGACAG	171
Db	1505	AACCTCGGAAACCGGAGCTAATACCGGATATATATTGAACCCGATGTTCAATAGTA	1448
QY	172	GGATCGAAGAAGCTTTCGCTTTTGGAGCGGCCGCAATGTCGATTAGCTAGTGGTGCGGTA	231
Db	1445	AAGACGGTTTTTGGCTGTCACTTATAGATGGAATCCGCGCGCATTTAGCTAGTTGGTAAGTA	1386
QY	232	AAGCCTTACCAAGGCGACGATCAGTATGTTGGTCTGAGAGACGACCAACCACTGGAC	291
Db	1385	ACGGCTTACCAAGGCAACGATGCGTAGCCGCACTGAGAGGGGTGATCGGCCACACTGGAAC	1328
QY	292	TGAGACACGCGCCGACATCTCTACGGGAGGCGACAGTGGGAAATTTTGGACAATGGCGCA	351
Db	1325	TGAGACACGCGTCCAGACTCTTACGGGAGGCGACAGTAGGGAAATCTTCGCAATGGCGCA	1266
QY	352	AGCGGATCCAGCAATGCGCGGTAGAGGAAGAAAGGACCTTCGGGTTGTAAAGCTCTTTGAG	411
Db	1285	AGCCTGACGAGCAACGCGCGGTAGAGTAAGTAAAGTCTTCGGAATCGTAAACTCTGTTAT	1206
QY	412	TCGAGAAAGAAAGGTTACGTTAAATTCGTGACTGACGCGTATCGACAGAAAGACA	471
Db	1205	TAGGGAAGAAACAATGTGTATAGTAAGTATGACAGCTTTGACGTTACCTAATCAGAAAGCC	1148
QY	472	CCGGCTAATCAGTTCGCGACGACGCGCGGTAACTAGTAGGGTGCAAGCTTAATGGAATT	531
Db	1145	ACGGCTAATCTACGTCCGACGACGCGCGGTAACTAGTAGGGTGCAAGCGTTATCCGGAATT	1088
QY	532	ACTGGGGGTAAAGGGTGCGACGCGCGCTTTGTAGTCAAGATGTAAATCCCGGGCTTAA	591
Db	1085	ATTGGGGGTAAAGCGCGGTAGGGCGGTTTTTAAAGTCTGATGTAAAGCCACGACTCA	1028
QY	592	CCTGGGAATTGCGTTTGAACAATCAAGGCTAGAGTGGCGAGGAGGAGTGGAAATTCAT	651
Db	1025	CGGTGAGGCGTCATTGGAACTGGAACAATTGAATGACGAAGAGAAAGTGAATTCAT	966
QY	652	GTGTAAGCATGAATTCGCTAGAGATATGGAAGAACATGATGCGGAGGACGCTCTCGG	711
Db	965	GTGTAGCGGTGAATATGCGCAGAGATATGAGAGAACACAGTGGGAGGCGCATCTTCTCGG	906
QY	712	GTTAACACTGACGCTCAATGACAGAAACGTGGGGAGCAACAGATTAAGATACCTGGTA	771
Db	905	TCTGTAACTGACGCTGATGTGCGAAACGTGGGGATCAACAGATTAAGATACCTGGTA	846
QY	772	GTGCACGCCCCAAGATGTCAACTAGTGT- - -GGGCTTATTAAGCTTGGTAAACGAG	829
Db	845	GTCCACGCCCCAAGATGATGCTAAGTGTAGGGGGTTTCCGCCCTTAAGTCTGAG	786
QY	830	CTAAACGCGTAGGTTGACCGCTTGGGAGTAGCGGTCCAGATTTAAACTCAAGAAATT	889
Db	785	CTTAACGCTTAAAGCACTCCGCTGGGAGTAGCACCCCAAGTTGAAACTCAAGAAATT	726
QY	890	GACGGGACCCGCAACAGCGGTGATATATGTGATTAATTCGATGCAACGGGAAAACT	949
Db	725	GACCGGGACCCGCAACAGCGGTGAGCATGTGTTTAATCGAAGCAACGGGAAACT	666
QY	950	TACCTACCCCTTGAATGAGGAATTTTCTAGAGATGATTAGTGC- - -TTCCGGAACGC	1006
Db	665	TACCAAACTTTGACATCTCTGACCCCTTAGAGATAGATTTTCCCTTCGGGGGACAG	606
QY	1007	TAAACAGAGTCTGATGCGTGTGCTGACGCTCGTGTCTGAGATGTTGGTTAAGTCCG	1066
Db	605	AGTACAGAGTGTGCATGTTGTGTCTCAGGCTCGTGTCTGAGATGTTGGTTAAGTCCG	546
QY	1067	CAACGAGCGCAACCTTGTCTATTATTTGCCATCTTTGGTTGGGCACTTTAATGAGACTG	1126
Db	545	CAACGAGCGCAACCTTAAAGTTGATGTTGCCATCTTAAGTTGGGCACTCTTAAGTTGACTG	486
QY	1127	CCGGTGAACAAACGGAGAAAGGTGGGATGACGCAAGTCCCTATGCGCTTAAGGGTGA	1186
Db	485	CCGGTGAACAAACCGAGAAAGGTGGGATGACGCAATCATCATGCCCCCTTAATGATTTGG	426
QY	1187	GGCTTTCACGTTAATCAATGGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGCTAA	1246

Accession	Sequence	Position
Db	GGCTACACACGTGCTACATGACACATACAAAGGATACGAAACCGCAGCGTCAAGCAAA	366
Qy	TCTCAGAAAGCGCGTCGTAGTCCGGATCGAGTCTGCAACTCGACTCCGTGAAGTCGAA	1306
Db	TCCCATTAAGTTGTTCTCAGTTCGGATTGTAGTCGAACCTCGATTATATAAAGTCGAA	306
Qy	TCGCTAGTAATCGCGGATCAGCATGCTCCGGGTAATACGTTCCCGGCTTTGTACACACC	1366
Db	TCGCTAGTAATCGTATGATCAGCATCTCAAGGTAATCGTTCCCGGCTTTGTACACACC	246
Qy	GCCGCTCACACACCGGAGTGGGTTTCAACAGAAACAGATGTCTAACCGTAAGGAGGAC	1426
Db	GCCGCTCACACCAAGATTTGTAAACCCGAAACCGGTGAGTAACCATTTGAGACTA	186
Qy	GCTTGCCACGGTGAATTCATGACTGGGGTG	1457
Db	GCCGTGGAAGTGGGACAAATGATTGGGGTG	155

```

RESULT 47
US-10-793-626-4460/c
; Sequence 4460, Application US/107933626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUJ480US
; CURRENT APPLICATION NUMBER: US-10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4460
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4460

```

Query Match	57.2%;	Score 833.8;	DB 10;	Length 3008;
Best Local Similarity	75.3%;	Pred. No. 3.1e-250;		
Matches 1108; Conservative	0;	Mismatches 347;	Indels 16;	Gaps 5;

QY 3 TGAACGCTGGCGGCAATGCTTTACATATCGAAGTCGCAAGCGGAGCATGCTTGCATCT 62  
Db 1525 TGAACGCTGGCGGCGCTGCTTAATATCAAGTCAGAGGAACAAGCGAGGCTTGCCTC 1466  
QY 63 GGTGGCG--AGTGGGGAGCGGGTGATGATGATGG--ACGTATCCGAAGAAGGAGGAGT 119  
Db 1465 TCTGACGTTAGGGGGAGCGGGTGAGTAACATGGAATTAACCTATTAAGACTGGGAT 1406  
QY 120 AACGATCGAAGATGTGCTAATACCGCAT-----TACTTAAGAGGAAGAAGCAG 171  
Db 1405 AACCTTCGGGAACCGGAGCTTAATACCGATATATATTGAACCGATGCTTCAATAGCA 1346  
QY 172 GGATCGAAGAACCTTGGCGCTTTGGAGCGGCCGATGTCTGATTAGTACTAGTTGGTGGGTA 231  
Db 1345 AAGACGGTTTCTGCTCACTTAATGATGATTCGGCGCGCANTTACTAGTTGGTAAAGTA 1286  
QY 232 AAGGCTTACCAAGCGGACGATCAGTAGTTGATCTGAGAGGACGACCACTGGGAC 291  
Db 1285 ACGGCTTACCAAGGCAAGCATGCGTAGCGGACCTGAGAGGGTGATTCGGCACACTGGAAAC 1286  
QY 292 TGAACAACGCGCCAGACTCTTAACGGAAGCAGACAGTGGGAAATTTTGGACAATGGCGCA 351  
Db 1225 TGAACAACGCTCAGACTCTTAACGGAAGCAGACAGTAGGAATCTTCCGCAATGGCGCA 1166  
QY 352 AGCTTGATCAAGCAATGCCGCGTAGTGAAGGAAGGCTTCCGGATTGTAAAGCTCTTCAAG 411

Db	1165	AGCTTGAACGAGCAACCCCGCTGATGATGAAGAGTCTTCGGATTCGTAAACACTCTGTAT	1106
Qy	412	TCGAGAAAGAAAGGTTACCGGTAAATATATCTGACTCATGACGGTATTCACAGAAAGCA	471
Db	1105	TAGGAGAAACAAATGTGTAAAGTAACTATCATCAAGTCTTGACGGTAACTTAATCAGAAAGCC	1046
Qy	472	CCGGCTAACTACGTGCCAGACGCCCGGCTAATACGTAGGGTGCAGAGGTTAAATTCGAAT	531
Db	1045	ACGGCTAACTACGTGCCAGACGCCCGGCTAATACGTAGGGTGCAGAGGTTATCCGAAAT	986
Qy	532	ACTAGGGCTAAAGGGTGCAGAGGGGCTTTGTAAGTACAGATGTGAATATCCCGGCTTAA	591
Db	985	ATTAGGGCTAAAGGGTGCAGAGGGGCTTTTAAAGTGTGAATGTGAAGCCACGGCTCAA	926
Qy	592	CCTGGGAATTCGCTTTGAAACTACAGGCTAAGATGTGGACAGAGAGGTGAATTCAT	651
Db	925	CCGTGAGAGGCTATTTGAAACTGGAAAACTTGATGTGACAGAAAGAGAAATTCAT	866
Qy	652	GTTGACAGTGAATTCGGTAGATATGGAAGAACATCGATTCGGAGAGGCAAGCTCCG	711
Db	865	GTTGACCGGTAAATTCGGTAGATATGGAAGAACACACAGTTCGGAGAGGCAAGCTTCG	806
Qy	712	GTTAAACCTGACGCTCAATGACAGAAACGGTGGGAGCAACAGATTAAGATACCTGTA	771
Db	805	TCTGTAACTGACGCTGATGTCCGAAGGTGGGATCAACAGATTAAGATACCTGTA	746
Qy	772	GTCCACGCGCTTAAACGATGTCAACTAGTGT--GGGCTTTATTAAGCTTGTGTAACAAG	829
Db	745	GTCCACGCGCTTAAACGATGTCTAAGTGTAAAGGGGTTTCCGCCCTTATGTCTCGAG	686
Qy	830	CTAACGCGTGAAGTTGACCGCTTGGGAGTACGGTCCGACATTTAAATCTCAAGAAAT	889
Db	685	CTAACGCTTAAAGCACTCCGCTGGGAGTTCGACCCGCAAGTGTGAATCTCAAGAAAT	626
Qy	890	GACGGGAGACCCGCAACAGCGGTGATTAATGTGAATTAATTCAGAGCAACGGAAAACT	949
Db	625	GACGGGAGACCCGCAACAGCGGTGAGCATGTGTATTAAATGTGAAGCAACGGAAAACT	566
Qy	950	TACCTAACCTTGAACATGTAGGAAATTTTCTAGATAGA---TTAGTCTTCGGAAAGC	1006
Db	565	TACCAATCTTGACATCTCTGATCCCTTGAGATTAAGGTTTCCCTTGAGGGAGACAG	506
Qy	1007	TAAACAGGTGTGCATGGAGGCTGTGTCAGCTGTGTGTGATGATGTTGGTTAAAGTCCG	1066
Db	505	AATGACAGGTGTGATGATGTGTGTGTCAGCTGTGTGTGATGATGTTGGTTAAAGTCCG	446
Qy	1067	CACGACGGCAACCTTGTCTATTAATGCCATCATTTGGTGGGCACTTAAATGAGACTG	1126
Db	445	CACGACGGCAACCTTAAAGTCTAGTGGCACTTAATGTTGGGCACTTAATGTTGACTG	386
Qy	1127	CGGTGACAAACCGGAGGAGGTGGGATGACGTCAAGTCTCTCATGGCCCTTAATGGTAG	1186
Db	385	CCGGTGAACAAACCGGAGGAGGTGGGATGACGTCAATCATGATGCCCTTAATGATTTG	326
Qy	1187	GCGTTCAACAGTAAATCAATGGCGGTAACAAGGGTTCCAACCCGCGAGGGAGCTTA	1246
Db	325	GCGTCAACAGGTAAATCAATGACAATCAAAAGGCAACGAAACCGCAGGCTCAAGCAA	266
Qy	1247	TCTCAGAAAGCCGCTGTAGTCCGATTCGGAGTCTTGCAACTCGATCCGTGAAGTCGGA	1306
Db	265	TCCCAATAAAGTGTCTCAGTTCGATTTGTGTGTGCAACTCGATTAATGAGCTGGAA	206
Qy	1307	TGCGTAGTAATCGCGATTCAGCATTCGCGGTAAATAGCTTCCCGGCTTGTACACAC	1366
Db	205	TGCGTAGTAATGTGTAGTCAAGATCTACCGGTAAATAGCTTCCCGGCTTGTACACAC	146
Qy	1367	GCCCGTCAACACCTGAGGTGGGTTTCAACAGAACAGTATGTCTAACCGTAAGAGGAC	1426
Db	145	GCCCGTCAACACAGAGGTTTGTAAACCCGAAGCCGCTGGAAGTAAACATTTGGAGCTA	86
Qy	1427	GCTTGCACGCTGAGATTCATGACTGGGGTG	1457
Db	85	GCGGTGAAAGGTGGCAAAATGATTTGGGGTG	55



## RESULT 48

```
US-10-793-626-3955/c
; Sequence 3955, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3955
; LENGTH: 3253
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; US-10-793-626-3955
```

```
Query Match      57.2%; Score 833.8; DB 10; Length 3253;
Beef Local Similarity 75.3%; Pred. No. 3.1e-250;
Matches 1108; Conservative 0; Mismatches 347; Indels 16; Gaps 5;
```

```
QY      3 TGAACGCTGCGCGCATGCTTTACATGCAAGTCGAAACGCGACAGCGATGCTTGATCT 62
DB      1476 TGAACGCTGCGCGCGCATGCTTTACATGCAAGTCGAAACAGAGAGAGAGCTTGCTTC 1417
QY      63 GGTGCGG--AGTGGCGGAGCGGCGTGAATGATCGG--AACGATACAGAAAGGGGGGT 119
DB      1416 TCTGACGTTAGCGGCGGAGCGGCGTGAATGATGATGATGATGATGATGATGATGATGAT 1357
QY      120 AACGATCGAAGATGCTTAATATACCGCATATCTCT-----AAGAGGAAAGCAG 171
DB      1356 AACCTCGGAAACCGGAGCTTAATACCGGATATATATTAACCGCATGCTTAATGATGAT 1297
QY      172 GATCGAAGAACCTTGCGCTTTTGAGCGCGCGATGCTGATTAAGTAACTTGTGCGGCTA 231
DB      1286 AAGACGGTGTGCTGCTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1237
QY      222 AAGGCTTACCAAGGCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 291
DB      1236 ACGGCTTACCAAGGCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1177
QY      292 TGAAGACAGGCGCGAGCTCTTAACGAGGAGCAGAGTGGGAAATTTTGAACAATGGGCGCA 351
DB      1176 TGAAGACAGGCTTCAACATCTCTTAACGAGGAGCAGAGTGGGAAATTTTCCGCAATGGGCGCA 1117
QY      352 AGCCTGATCAGCAATGCGCGCTGATGAAGAAGGCTTGCGGTTGTAAGCTCTTTCAG 411
DB      1116 AGCCTGAGGAGCAACGCGCGCTGATGAAGAAGGCTTGCGGTTGTAAGCTCTTTCAT 1057
QY      412 TCGAAGAGAAAGGTTACGTTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 471
DB      1056 TAGGAGAAAGAAATGTTGTTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 997
QY      472 CCGGCTTACATGCTGCGACAGCGCGGATTAATACGTAAGGAGTCAAGCGCTTAATCGGAAT 531
DB      996 ACGGCTTACATGCTGCGACAGCGCGGATTAATACGTAAGGAGTCAAGCGCTTAATCGGAAT 937
QY      532 ACTGGCGTAAAGGCTGCGACGCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 591
DB      936 AATTGGCGTAAAGGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 877
QY      592 CTTGGGAAATGCGTTTGAATCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 651
DB      876 CCGTGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 817
```

```
QY      652 GTTAGCAGTGAATGCTAGATATGAAAGAACATGATGCGAGGAGCGCTCTG 711
DB      816 GTTAGCAGTGAATGCTAGATATGAAAGAACATGATGCGAGGAGCGCTCTG 757
QY      712 GTTAACACTGACGCTGATGACAGAAAGCGTGGGAGCAACAGATTTGATACCTGGTA 771
DB      756 TCTGTACTGACGCTGATGACAGAAAGCGTGGGAGCAACAGATTTGATACCTGGTA 697
QY      772 GTTCAAGCGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829
DB      636 GTTCAAGCGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 637
QY      830 CTAAAGCGTGAAGTTGACCGCTGCGGAGTACGCTGCGAAGTTAACTCAAGGAAT 889
DB      636 CTAAAGCGTGAAGTTGACCGCTGCGGAGTACGCTGCGAAGTTAACTCAAGGAAT 577
QY      890 GACGGGAGCCGCGACAGCGGCTGATTTATGATGATTTATGATGATGATGATGATGATGATGAT 949
DB      576 GACGGGAGCCGCGACAGCGGCTGATTTATGATGATTTATGATGATGATGATGATGATGATGAT 517
QY      950 TACCTAACCTTGACATGAGGAATTTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 1006
DB      516 TACCTAACCTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 457
QY      1007 TAAACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066
DB      456 AGTACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 397
QY      1067 CAACGAGCGCAACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126
DB      396 CAACGAGCGCAACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 337
QY      1127 CCGGTGAAGAACCGGAGAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1186
DB      336 CCGGTGAAGAACCGGAGAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 277
QY      1187 GAGCTTCAACAGTAAATACATGCGCGCTGACAGAGGTTGCCAACCGCGAGGAGGAGCTTAA 1246
DB      276 GAGCTTCAACAGTAAATACATGCGCGCTGACAGAGGTTGCCAACCGCGAGGAGGAGCTTAA 217
QY      1247 TCTCAAGAAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306
DB      216 TCCCATTAAGTTGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 157
QY      1307 TCGCTAGTAATCGGGGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
DB      156 TCGCTAGTAATCGGGGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 97
QY      1367 GCCGCTCACACATGAGAGTGGGTTTCAACAGAGAGTGAAGTGAACCGTAAAGAGGGC 1426
DB      96 GCCGCTCACACATGAGAGTGGGTTTCAACAGAGAGTGAAGTGAACCGTAAAGAGGGC 37
QY      1427 GCTTGCACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
DB      36 GCCGCTCACACATGAGAGTGGGTTTCAACAGAGAGTGAAGTGAACCGTAAAGAGGGC 6
```

## RESULT 49

```
US-10-793-626-3967/c
; Sequence 3967, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3967
; LENGTH: 3821
```



TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3967

Query Match 57.2%; Score 833.8; DB 10; Length 3821;  
Best Local Similarity 75.3%; Pred. No. 3.3e-250;  
Matches 1108; Conservative 0; Mismatches 347; Indels 16; Gaps 5;

QY 3 TGAACGCTGGCGGCTTACATGCAAGTCGAAAGCGACACCGATGCTTGCATCT 62  
DB 3624 TGAACGCTGGCGGCTTACATGCAAGTCGAAAGCGACACCGATGCTTGCATCT 3565  
QY 63 GGTGGCG--AGTGGGGAACGGGTGATGATGCGG--AAGCATTCGAAGAAGGGGGGT 119  
DB 3564 TCTGACGTTAGCGGGAACGGGTGATGATGCGG--AAGCATTCGAAGAAGGGGGGT 3505  
QY 120 AACGATGGAAGATGTGCTAATACCGCATF-----TACTTAAAGAGGAAGCAAG 171  
DB 3504 AACCTGGGAACCGAGCTTAATACCGGATTAATTAATGAACCGCATGCTTCAATAGTA 3445  
QY 172 GGAATGGAAGACCTTGGCTTTTGAAGCGCGCATGTCTGATTAAGCTAGTTGGTGGGTA 231  
DB 3444 AAGAGGTTTGTCTGTCACTTAATGATGATCCGGCCGATTAAGCTAGTTGGTAAAGTA 3385  
QY 232 AAGGCTTACCAAGCGACATCAGTATGTTGTTCTGAGAGAGACGACGACACTGGGAC 291  
DB 3384 ACGGCTTACCAAGCGACATCAGTATGTTGTTCTGAGAGAGACGACGACACTGGGAC 3325  
QY 292 TGAAGACGCGCCAGACTCTCTACGAGGAGCAGACGAGGGAATTTTGAACATGGGCGCA 351  
DB 3324 TGAAGACGCGCTCAAGCTCTCTACGAGGAGCAGACGAGGGAATTTTCCGCAATGGGCGAA 3265  
QY 352 AAGCTGATCAGCAATATGCGGCTGATGAGGAAGAGGCTTCCGCTTGAAGCTCTTTCAG 411  
DB 3264 AAGCTGATCAGCAATATGCGGCTGATGAGGAAGAGGCTTCCGCTTGAAGCTCTTTCAT 3205  
QY 412 TGAAGAGGAAGGTTACGCTAATATGCTGATCATGATCGGATTCGACAGGAAGACA 471  
DB 3204 TGAAGAGGAAGCAATATGCTAATATGCTGATCATGATCGGATTCGATCAAGGAAGCC 3145  
QY 472 CCGGCTTACCTAGCTGTCAGACGCGGCTTAATAGTATGAGGCTGCAACGTTATCGAAT 531  
DB 3144 ACGGCTTACCTAGCTGTCAGACGCGGCTTAATAGTATGAGGCTGCAACGTTATCGGAAT 3085  
QY 532 ACTGGCGTAAAGGCTGCGAGCGGCTTGTAAAGTCAATGTAATCCCGGCTTAA 591  
DB 3084 ATTGGGCGTAAAGCGGCGGCTTGTAAAGTCAATGTAATCCCGGCTTAA 3025  
QY 592 CCGGGAATGCTTGAACCTCAAGGCTAAGTGTGCGAGAGGAGGTGGAATTCAT 651  
DB 3024 CCGTGAAGGCTCAATGGAACCTCAAGGCTAAGTGTGCGAGAGGAGGTGGAATTCAT 2965  
QY 652 GTGTACAGAGGAATGTGATGATTAAGGAACATCATGAGGAGGAGGAGGCTCTGG 711  
DB 2964 GTGTACAGAGGAATGTGATGATTAAGGAACATCATGAGGAGGAGGAGGAGGCTCTGG 2905  
QY 712 GTTAAACATGACGCTCATGACGAAAGCTGGGAGCAACAGATTAATGATACCTGTA 771  
DB 2904 TCTGTAACTGACGCTCATGATGCGAAAGCTGGGAGTCAACAGATTAATGATACCTGTA 2845  
QY 772 GTTCAAGCCTTAAACGATGCTAATGTTT--GGGCTTATTAAGCTTGTGTAAGAG 829  
DB 2844 GTTCAAGCCTTAAACGATGCTAATGTTT--GGGCTTATTAAGCTTGTGTAAGAG 2785  
QY 830 CTAAAGGCTGAAGTGAACGCTGGGAGTAAAGCTGCAAGATTAATTAAGCAAGAT 889  
DB 2784 CTAAAGGCTGAAGTGAACGCTGGGAGTAAAGCTGCAAGATTAATTAAGCAAGAT 2725  
QY 890 GACGGGAGCCGCAACAGCGGTGATTAATGATTAATTCATGCAAGCGGAAGAACT 949

DB 2724 GACGGGAGCCGCAACAGCGGTGAGCATGTGTTAATTCGAACACCGGAAGAACT 2665  
QY 950 TACCTACCTTGAATGATGCGAATTTCTAGAGATGATAGTGC---TTCGGGAACGC 1006  
DB 2664 TACCAATCTTGAATGATGCGAATTTCTAGAGATGATAGTGC---TTCGGGAACGC 2605  
QY 1007 TAAACAGGCTGATGATGCGAATTTCTAGAGATGATAGTGC---TTCGGGAACGC 1066  
DB 2604 AGTACAGGCTGATGATGCGAATTTCTAGAGATGATAGTGC---TTCGGGAACGC 2545  
QY 1067 CAACAGGCTGATGATGCGAATTTCTAGAGATGATAGTGC---TTCGGGAACGC 1126  
DB 2544 CAACAGGCTGATGATGCGAATTTCTAGAGATGATAGTGC---TTCGGGAACGC 2485  
QY 1127 CCGGTGCAACCGGAGAGTGGGATGACGTCAGTCTCATGCGCTTATAGGCTAG 1186  
DB 2484 CCGGTGCAACCGGAGAGTGGGATGACGTCAGTCTCATGCGCTTATAGGCTAG 2425  
QY 1187 GCGTTCACATGATTAATGATGCGGCTGACAGAGGTTGCCAACCGCGAGGGGAGCTTA 1246  
DB 2424 GCGTTCACATGATTAATGATGCGGCTGACAGAGGTTGCCAACCGCGAGGGGAGCTTA 1246  
QY 1247 TCTGAAAGGCGCTGATGCGGCTGACAGTCTGCAACTGCACTCCGTGAAGTGGAA 1306  
DB 2364 TCCATTAAGTGTCTTCAATGCGGCTGACAGTCTGCAACTGCACTCCGTGAAGTGGAA 2305  
QY 1307 TCGCTAATATGCGGATCAGATGTCGCGGTAATGCTTCCGGCTTGTACACAC 1366  
DB 2304 TCGCTAATATGCGGATCAGATGTCGCGGTAATGCTTCCGGCTTGTACACAC 2245  
QY 1367 GCCGCTCACATGAGGAGTGGTTCACCAAGACAGTGTCTAACCGTAAGAGGAC 1426  
DB 2244 GCCGCTCACATGAGGAGTGGTTCACCAAGACAGTGTCTAACCGTAAGAGGAC 2185  
QY 1427 GCTTCCACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
DB 2184 GCTTCCACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2154

RESULT 50  
US-11-273-617-9  
Sequence 9, Application US/11273617  
Publication No. US20060067924A1  
GENERAL INFORMATION:  
APPLICANT: Lee, Margie D.  
APPLICANT: Hofacre, Charles  
APPLICANT: Harmon, Barry  
TITLE OF INVENTION: Probiotic Bacteria and Methods  
FILE REFERENCE: 18-03  
CURRENT APPLICATION NUMBER: US/11/273,617  
CURRENT FILING DATE: 2005-11-14  
PRIOR APPLICATION NUMBER: US 60/470,807  
PRIOR FILING DATE: 2003-05-14  
PRIOR APPLICATION NUMBER: WO us2004/015378  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Patentin version 3.2  
SEQ ID NO 9  
LENGTH: 1510  
TYPE: DNA  
ORGANISM: Lactobacillus aviarius  
US-11-273-617-9

Query Match 57.1%; Score 832.4; DB 18; Length 1510;  
Best Local Similarity 76.6%; Pred. No. 7.1e-250;  
Matches 1074; Conservative 0; Mismatches 316; Indels 12; Gaps 4;

QY 68 CGAGTGGCGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 126  
DB 105 CGAGTGGCGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 164  
QY 127 CGAAGATGTCTAATACCGCATATATCTTAAGAGGAAGAGGAGATGAAAGAC--- 183

Db 165 GGAACCAATGCTAATACCGTATTAACGATGACCGCATGCTCATTTAATGTAAGGTGG 224  
Qy 184 -----CTTGCGCTTTTGGAGCGCGCATGCTGATTAAGTATGCTGGGGTAAAGGCTT 238  
Db 225 TTTTGCCTATCGCTTTTGGATGAGCCCGCGCTATTAATTAAGTGGGGTAAAGGCTT 284  
Qy 229 ACCAAGGAGAGATCAGTATGTTGATCTGAGAGAGACGACGACGACGACGACGACGACGAC 298  
Db 285 ACCAAGGAGATATACGTAAGCCGATTTGAGAGACTGATGCGCCAAATGAGACTAGACA 344  
Qy 299 CGGCGCAGACTCTTAACGAGAGGACGAGTGGGAAATTTTGAACAATGGCGCGCAACCTTGA 358  
Db 345 CGGCGCATATCTCTTAACGAGAGGACGAGTGGGAAATTTTGAACAATGGCGCGCAACCTTGA 404  
Qy 359 TCCAGCAATGCCGCTGATGTAAGAAAGGCTTGGGTTGTAAGCTCTTTCACTGAGAA 418  
Db 405 TGGAGCAACGCGCGCTGATGTAAGAAAGGCTTGGGTTGTAAGCTCTTTCACTGAGAA 464  
Qy 419 GAAAGGTTAGCGTAAATTAATGCTGATCATGACGGTATCGACAGAGAGACCGGCTA 478  
Db 465 GAAATTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 524  
Qy 479 ACTACGTCGACGACGCGCGGTATACGTAAGGAGTCAAGCGTAAATCGAAATTAATTAATTA 538  
Db 525 ACTACGTCGACGACGCGCGGTATACGTAAGGAGTCAAGCGTAAATCGAAATTAATTAATTA 584  
Qy 539 GTAAAGGTCGCGCAGCGGCTTTGTAATGATGTAAGTGTAAATCCCGGGCTTAATTAATTA 598  
Db 585 GTAAAGGTCGCGCAGCGGCTTTGTAATGATGTAAGTGTAAATCCCGGGCTTAATTAATTA 644  
Qy 599 ATTGCGTTTAAACTACAGAGGCTAGAGTGGCAGAGAGGAGGAGTATTCATGCTGAGC 658  
Db 645 TGTGATTAAGTAACTGAGAGGCTAGAGTGGCAGAGAGGAGTATTCATGCTGAGC 704  
Qy 659 AGTGAATGCGTGAAGATATGAGAGACATGATGCGGAGGACGCTTGGGTTAAACA 718  
Db 705 GGTGAATGCGTGAAGATATGAGAGACATGATGCGGAGGACGCTTGGGTTAAACA 764  
Qy 719 CTGAGCGTCTATGCAAGAAAGCGTGGGAGCAAAAGATTAATACCTTGGTATGCTCAGC 778  
Db 765 CTGAGCGTCTATGCAAGAAAGCGTGGGAGCAAAAGATTAATACCTTGGTATGCTCAGC 824  
Qy 779 CCTTAAGGATGCAATAGTGTGGGCTTAATTAAGGCTTGGGTTAAACA 836  
Db 825 CCTTAAGGATGCAATAGTGTGGGCTTAATTAAGGCTTGGGTTAAACA 884  
Qy 837 GTGAAGTGAACCGCTGGGAGTACGATGCGCAAGATTAAACTCAAGAAATTAAGTGA 896  
Db 885 AATAAGCATTCGCGCTGGGAGTACGATGCGCAAGATTAAACTCAAGAAATTAAGTGA 944  
Qy 897 ACCGCGACAGCGGCTGATTAATGATGCAATGCAAGCGCAAAACCTTAATTA 956  
Db 945 GCCCGCACAGCGGCTGATTAATGATGCAATGCAAGCGCAAAACCTTAATTA 1004  
Qy 957 CCTTAAGGATGCAATAGTGTGGGCTTAATTAAGGCTTGGGTTAAACA 1015  
Db 1005 TCTTAAGGATGCAATAGTGTGGGCTTAATTAAGGCTTGGGTTAAACA 1064  
Qy 1016 TGTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1075  
Db 1065 TGTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1124  
Qy 1076 CAACCTTGTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1135  
Db 1125 CAACCTTGTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1184  
Qy 1136 AACCGAGAGAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1195  
Db 1185 AACCGAGAGAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1244  
Qy 1196 CGTATATCAATGCGCGCTACAGAGGTTGCCAACCGCGAGGAGGAGCTTAATCTCAGAAA 1255  
Db 1245 CGTATATCAATGCGCGCTACAGAGGTTGCCAACCGCGAGGAGGAGCTTAATCTCAGAAA 1304

Qy 1256 GCGGCTGTAGTCCGAGTCCGAGTCTGCAACTCGATCCGTAAGTCCGAATCCCTAGTA 1315  
Db 1305 GTGCTTCTCACTTGGATGAGGCTGCACTCCCTGTCATGAAGTCCGTAAGTCCCTAGTA 1364  
Qy 1316 ATCCGAGTCAAGATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACCGCGCTGAC 1375  
Db 1365 ATCCGAGTCAAGATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACCGCGCTGAC 1424  
Qy 1376 ACCATGGAGTGGGTTTCAACCAAGCAGTATCTTAACGTAAGAGGCGCTTGGCAC 1435  
Db 1425 ACCATGGAGTGGGTTTCAACCAAGCAGTATCTTAACGTAAGAGGCGCTTGGCAC 1484  
Qy 1436 GGTGATTAATCACTGCGGCTG 1457  
Db 1485 GGTGATTAATCACTGCGGCTG 1506

RESULT 51  
US-10-953-392-1  
; Sequence 1, Application US/10953392  
; Publication No. US2005028967A1  
; GENERAL INFORMATION:  
; APPLICANT: CAMBIA  
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells  
; FILE REFERENCE: CAMBIA 4148  
; CURRENT FILING DATE: 2004-09-28  
; PRIOR FILING DATE: 2004-06-28  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 1408  
; TYPE: DNA  
; ORGANISM: Rhizobium spp. NGR234 (strain ANU240)  
US-10-953-392-1

Query Match 57.1%; Score 832; DB 10; Length 1408;  
Best Local Similarity 77.8%; Pred. No. 9.3e-250;  
Matches 1087; Conservative 0; Mismatches 265; Indels 46; Gaps 5;

Qy 66 GCGAGTGGCGGAGCGGCTGATTAATGATCGAACTGATCCAGAAAGGCGGTTAAAGCA 125  
Db 31 GCGAGTGGCGGAGCGGCTGATTAATGATCGAACTGATCCAGAAAGGCGGTTAAAGCA 90  
Qy 126 TCGAAGATGTCTAATACCGCATATACCTTAAGAGAGAAAGCAGGAGATCGAAAGACT 185  
Db 91 GCGAAGATGTCTAATACCGCATATACCTTAAGAGAGAAAGCAGGAGATCGAAAGACT 145  
Qy 186 TGGCGTTTGGAGCGCGCATGCTGATTAAGTATGCTGATGCTGATGCTGATGCTGATGCTGAT 245  
Db 146 -----GGATGAGCGCGCGTGGATTAAGTATGCTGATGCTGATGCTGATGCTGATGCTGAT 196  
Qy 246 CGAGATCAAGTATGCTGATGATGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 305  
Db 197 CGAGATCAAGTATGCTGATGATGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 256  
Qy 306 GACTCTACGAGGAGCAGCAGTGGGAAATTTTGAACAATGGCGCAAGCTGATCCAGCA 365  
Db 257 AACTCTACGAGGAGCAGCAGTGGGAAATTTTGAACAATGGCGCAAGCTGATCCAGCA 316  
Qy 366 ATGCGCGTGAAGTGAAGAGCCCTCGGTTGTAAGCTCTTCAAGTGAAGAGAGAGAGAGAG 425  
Db 317 ATGCGCGTGAAGTGAAGAGCCCTCGGTTGTAAGCTCTTCAAGTGAAGAGAGAGAGAGAG 371  
Qy 426 TTAAGTAAATTAATGCTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 485  
Db 372 -----TTAAGTAAATTAATGCTGATCATGATGATGATGATGATGATGATGATGATGATGAT 411  
Qy 486 GCGAGCAGCGCGGTTAATGATGAGTGAAGCGTTAATGGAATTAATGGAATTAATGGAATTAATG 545  
Db 412 GCGAGCAGCGCGGTTAATGATGAGTGAAGCGTTAATGGAATTAATGGAATTAATGGAATTAATG 471

546 GTGCGCAGCGCGCTTTGTAAGTCAGATGTGAATCCCGGCGCTTAACCTGGGAATTGGCT 605  
DB GCACCTAGGCGGACATTTAAAGTCAGGGGTGAATCCCGGGGCTCAACCCCGGACCTGCT 531  
606 TTGAAACTCAAGGCTAAGGTGGCAGAGGGAGGTGAATTCATGTCTTACAGTGA 665  
DB TTGATATCTGGGTGTCTAGAGTCCGGAGAGGTGAAGTAATCCGAGTGTGAAGGTGA 591  
666 TGGCGTAGATATGGAAGACATCGATGCGAAGGACGCTCTCGGGTAAACATGACGC 725  
DB TTGATATCTGGGTGTCTAGAGTCCGGAGAGGTGAAGTAATCCGAGTGTGAAGGTGA 591  
552 TTGATATCTGGGTGTCTAGAGTCCGGAGAGGTGAAGTAATCCGAGTGTGAAGGTGA 651  
726 TCATGCAAGAAAGCGTGGGAGCAACAGAGTTAGTACCTGTGATGTCACGCGCTTAA 785  
DB TGAAGTGCAGAAAGCGTGGGAGCAACAGAGTTAGTACCTGTGATGTCACGCGCTTAA 711  
786 CGATGTCACTAGTGTGGGCTTTATAGGCTTGTGAAGGCTTAAGCGGCTGAAGTTG 845  
DB CGATGTCACTAGTGTGGGCTTTATAGGCTTGTGAAGGCTTAAGCGGCTGAAGTTG 771  
712 CGATGTCACTAGTGTGGGCTTTATAGGCTTGTGAAGGCTTAAGCGGCTGAAGTTG 771  
846 ACCGCGTGGGAGTACGCTCGCAAGATTAACCTCAAGAAATTGACGGGAGCCGAC 905  
DB TCCGCTGGGAGTACGCTCGCAAGATTAACCTCAAGAAATTGACGGGAGCCGAC 831  
906 ACCGCGTGAATATGTGATTAATTCGATGCAACGCGAAACCTTAACCTGATCAAT 965  
DB AGCGGTGAGCAATGTGTTTAATTCGAGCAACGCGCAACCTTAACCTGATCAAT 891  
832 AGCGGTGAGCAATGTGTTTAATTCGAGCAACGCGCAACCTTAACCTGATCAAT 891  
966 GTAGCAATTTTCTAGAG-----TAGATTAAGTCTTGGGAAAGCTTAACAGGTGCTG 1020  
DB CCCGCTGGGAGTACGAGATGCTATCTTCAGTTCGCTGAGCGGAGACAGGTGCTG 951  
892 CCCGCTGGGAGTACGAGATGCTATCTTCAGTTCGCTGAGCGGAGACAGGTGCTG 951  
1021 CATGCTGTCTGAGCTCGCTGTCGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1080  
DB CATGCTGTCTGAGCTCGCTGTCGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1011  
952 CATGCTGTCTGAGCTCGCTGTCGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1011  
1081 CTTGTCAATTAATTCGATCAATTTGTTGGGCACTTTAATGAGACTCCGGTGAACAAACG 1140  
DB CTCGCCCTTAAGTTGCCAGATTTGGTTGGGCACTTAAGGGAGCTCCGGTGAATAGCCG 1071  
1012 CTCGCCCTTAAGTTGCCAGATTTGGTTGGGCACTTAAGGGAGCTCCGGTGAATAGCCG 1071  
1141 -GAGGAAGTGGGAGTGAAGCTCAAGTCTCATAGCCCTTAATGGGTAGGGCTTCAACGTA 1199  
DB AGAGGAAGTGGGAGTGAAGCTCAAGTCTCATAGCCCTTAATGGGTAGGGCTTCAACGTA 1131  
1072 AGAGGAAGTGGGAGTGAAGCTCAAGTCTCATAGCCCTTAATGGGTAGGGCTTCAACGTA 1131  
1200 ATTCATATGCGCGCTTACAGAGGGTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGC 1259  
DB CTACCAATGTTGTGACAGTGGGAGCGAGACCGCGAGGGTGAAGTCTTCAAAAGC-C 1190  
1132 CTACCAATGTTGTGACAGTGGGAGCGAGACCGCGAGGGTGAAGTCTTCAAAAGC-C 1190  
1260 GTGCTAGTCCGAGTCCGAGTCTGCAACTCGACTCCGTGAAGTCCGATCTAGTATCG 1319  
DB ATTCAGTGTGAGTGTGACACTCTCCAACTCGAGTGAAGTGAAGTGGCTAGTATCG 1250  
1191 ATTCAGTGTGAGTGTGACACTCTCCAACTCGAGTGAAGTGAAGTGGCTAGTATCG 1250  
1320 CGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTTGACACACCGCCGTCACCA 1379  
DB CAGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTTGACACACCGCCGTCACCA 1310  
1251 CAGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTTGACACACCGCCGTCACCA 1310  
1380 TGGGATGTTGTTTCAACAGAAAGCAGAGTGAAGTGAAGGAGGGGCTTGGCAGGTTG 1439  
DB TGGGATGTTGTTTCAACAGAAAGTGAAGTGAAGTGAAGGAGGGGCTTGGCAGGTTG 1370  
1311 TGGGATGTTGTTTCAACAGAAAGTGAAGTGAAGTGAAGGAGGGGCTTGGCAGGTTG 1370  
1440 AGATTCACTGAGTGGGTTG 1457  
DB GGCTCAGCGACTGGGGTGG 1388

## RESULT 52

US-10-954-147-1  
; Sequence 1, Application US/10954147  
; Publication No. US20050289672A1  
; GENERAL INFORMATION:  
; APPLICANT: CAMBIA  
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells

FILE REFERENCE: CAMBIA 414  
; CURRENT APPLICATION NUMBER: US/10/954,147  
; CURRENT FILING DATE: 2004-09-28  
; PRIOR APPLICATION NUMBER: 60/583,426  
; PRIOR FILING DATE: 2004-06-28  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 1  
; LENGTH: 1408  
; TYPE: DNA  
; ORGANISM: Rhizobium spp. NGR234 (strain ANU240)  
US-10-954-147-1  
Query Match 57.1%; Score 832; DB 10; Length 1408;  
Best Local Similarity 77.8%; Pred. No. 9.3e-250;  
Matches 1087; Conservative 0; Mismatches 265; Indels 46; Gaps 5;  
66 GCGAGTGGCGGACGGGTGATGATTCGAAACGATTCGAAAGAGGGGTTAAGCA 125  
DB GCGAGTGGCGGACGGGTGATGATTCGAAACGATTCGAAAGAGGGGTTAAGCA 125  
31 GGGAGCGGACAGCGGTGATGATTCGAAACGATTCGAAAGAGGGGTTAAGCA 90  
DB GGGAGCGGACAGCGGTGATGATTCGAAACGATTCGAAAGAGGGGTTAAGCA 90  
126 TCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATCGAAAGACT 185  
DB GCGAAACTGTCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATCGAAAGACT 145  
91 GCGAAACTGTCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATCGAAAGACT 145  
186 TCCGCTTTTGGAGCGCGCGCATGCTGATTAAGTCTGTTGGGTTAAGGCTTACCAAG 245  
DB TCCGCTTTTGGAGCGCGCGCATGCTGATTAAGTCTGTTGGGTTAAGGCTTACCAAG 196  
146 -----GGATGAGCCCGCGTGTGATTAAGTCTGTTGGGTTAAGGCTTACCAAG 196  
246 CGACGATCAGTATGTTGTTGTTGAGAGAGCAGCAGCAGCACTGGGACTGAGACGCGCCA 305  
DB CGACGATCAGTATGTTGTTGTTGAGAGAGCAGCAGCAGCACTGGGACTGAGACGCGCCA 305  
197 CGACGATCAGTATGTTGTTGTTGAGAGAGCAGCAGCAGCACTGGGACTGAGACGCGCCA 256  
DB CGACGATCAGTATGTTGTTGTTGAGAGAGCAGCAGCAGCACTGGGACTGAGACGCGCCA 316  
306 GACTCTTACGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 365  
DB GACTCTTACGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 256  
257 AACTCTTACGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 316  
DB AACTCTTACGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 316  
366 ATGCGCGGTGATGAGAAAGGCTTCCGCTTGAAGCTCTTCACTGAGAAAGAAAG 425  
DB ATGCGCGGTGATGAGAAAGGCTTCCGCTTGAAGCTCTTCACTGAGAAAGAAAG 371  
317 ATGCGCGGTGATGAGAAAGGCTTCCGCTTGAAGCTCTTCACTGAGAAAGAAAG 371  
426 TTAAGTAAATATGTTGATCTGATGCGGTATCGACAGAGAAAGCAGCGCTTAACTACGT 485  
DB TTAAGTAAATATGTTGATCTGATGCGGTATCGACAGAGAAAGCAGCGCTTAACTACGT 372  
372 -----TAAATGCGGTAACTGAGAGAAAGCAGCGCTTAACTACGT 411  
486 GCCAGCAGCGCGGTAACTGATGAGGTGCAAGCTTAACTGAGAGAAAGCAGCGCTTAACT 545  
DB GCCAGCAGCGCGGTAACTGATGAGGTGCAAGCTTAACTGAGAGAAAGCAGCGCTTAACT 471  
412 GCCAGCAGCGCGGTAACTGATGAGGTGCAAGCTTAACTGAGAGAAAGCAGCGCTTAACT 471  
546 GTGCGGAGCGCGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 605  
DB GTGCGGAGCGCGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531  
472 GCACGTAGGCGGACATTTAAGTCAAGGGGTGAATCCCGGGGCTCAACCCCGGAACTGCGCT 531  
606 TTGAAACTCAAGGCTAAGGTGGCAGAGGGAGGTGAATTCATGTCTTACAGTGA 665  
DB TTGAAACTCAAGGCTAAGGTGGCAGAGGGAGGTGAATTCATGTCTTACAGTGA 591  
532 TTGATATCTGGGTGTCTAGAGTCCGGAGAGGTGAAGTAATCCGAGTGTGAAGGTGA 651  
666 TGGCGTAGATATGGAAGACATCGATGCGAAGGACGCTCTCGGGTAAACATGACGC 725  
DB TTGATATCTGGGTGTCTAGAGTCCGGAGAGGTGAAGTAATCCGAGTGTGAAGGTGA 651  
592 TTGATATCTGGGTGTCTAGAGTCCGGAGAGGTGAAGTAATCCGAGTGTGAAGGTGA 651  
726 TCATGCAAGAAAGCGTGGGAGCAACAGAGTTAGTACCTGTGATGTCACGCGCTTAA 785  
DB TGAAGTGCAGAAAGCGTGGGAGCAACAGAGTTAGTACCTGTGATGTCACGCGCTTAA 711  
786 CGATGTCACTAGTGTGGGCTTTATAGGCTTGTGAAGGCTTAAGCGGCTGAAGTTG 845  
DB CGATGTCACTAGTGTGGGCTTTATAGGCTTGTGAAGGCTTAAGCGGCTGAAGTTG 771  
712 CGATGTCACTAGTGTGGGCTTTATAGGCTTGTGAAGGCTTAAGCGGCTGAAGTTG 771  
846 ACCGCGTGGGAGTACGCTCGCAAGATTAACCTCAAGAAATTGACGGGAGCCGAC 905  
DB TCCGCTGGGAGTACGCTCGCAAGATTAACCTCAAGAAATTGACGGGAGCCGAC 831

Qy	906	AGCGGTGGATTATGTGGATTAAATTCGATGCCAAGCGGAAAAACCTTAACCTACCTTGGACAT	965
Db	832	AGCGGTGGAGCATGTGGTTAAATTCGAMACCAACGCGCAACCTTAACAGGCTTGAACAT	891
Qy	966	GTAGCGAAATTTTCTAGAGA-----TAGATTAGTCTTCGGGAACGGTAAACAGGTGCTG	1022
Db	892	CCCCGTCCGGGATACGAGAGATCGTATCTTTCAGTTCCGGCTGGACCGGAGACAGGTGCTG	951
Qy	1021	CATGCTGTCTGTCAGCTGTCGTGATGAGATGTTGGTTAAGTCCCGCAACAGCGCAAC	1088
Db	952	CATGCTGTCTGTCAGCTGTCGTGATGAGATGTTGGTTAAGTCCCGCAACAGCGCAAC	1011
Qy	1081	CTTGTGATTAATTGCCATCATTTGGTTGGGCACTTTAATGAGACTGCCGGTGAACAAACCG	1144
Db	1012	CTCGCCCTTAAAGTTGCCAGCATTTGGTTGGGCACTTAAAGGAGATGCGCGGTAAAGCCG	1077
Qy	1141	-GAGGAAGGTGGGGAATGACGTCAAGCTCTCATGACCCCTTAATGGGTAGGGCTTCAACGTA	1199
Db	1072	AGAGGAAGTGGGGAATGACGTCAAGCTCTCATGACCCCTTAATGGGCTGGGCTTCAACGTA	1133
Qy	1200	ATACAAATGCGCGCTACAGAGGGTTGCGAACCGCGAGGGGAGACTAATCTCAGAAAACCG	1255
Db	1132	CTACAAATGGTGTGACAGTGGGCAAGGAGACGCGAGGTCTGAGCTAATCTCAAAAAGC-C	1190
Qy	1260	GTCTGATGTCGGGATCGGAGATCTTGCAACTGCACTCCGTGAAGTCCGAAATGGCTGATTAATG	1311
Db	1191	ATCTCAATTCGGATTTGACATCTTGCAACTGAGTGCATGAAGTTGAATGCTTAAGTATATG	1255
Qy	1320	CGGATCAGCATGTCGCGGTGAATTAAGTTCGCCGGGTCTTGTACACACCGCCCGTCAACCA	1372
Db	1251	CAGATCAGCATGTCGCGGTGAATTAAGTTCGCCGGGTCTTGTACACACCGCCCGTCAACCA	1311
Qy	1380	TGGGAGTGGGTTTCAACCGAAGCAGGTAATCTTAACCGTAAGGAGGGCGCTTGCCACGGTG	1433
Db	1311	TGGGAGTGGGTTTCTACCCGAAAGTATGTCGCTTAACCGAAGGAGGAGCACTTAACACGGTA	1370
Qy	1440	AGATTCATGACATCGGGGTG 1457	
Db	1371	GGGTCAAGCACTGGGGTG 1388	
RESULT 53			
US-10-953-392-15			
; Sequence 15, Application US/10953392			
; Publication No. US20050289667A1			
; GENERAL INFORMATION:			
; APPLICANT: CAMBA			
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells			
; FILE REFERENCE: CAMBIA 414B			
; CURRENT APPLICATION NUMBER: US/10/953,392			
; CURRENT FILING DATE: 2004-09-28			
; PRIOR APPLICATION NUMBER: US 60/583,426			
; PRIOR FILING DATE: 2004-06-28			
; NUMBER OF SEQ ID NOS: 83			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 15			
; LENGTH: 1433			
; TYPE: DNA			
; ORGANISM: A. tumefaciens EHA105			
US-10-953-392-15			

Query Match	57.1%;	Score 832;	DB 10;	Length 1433;
Best Local Similarity	77.8%;	Pred. No. 9,4e-250;		
Matches 1087; Conservative	0;	Mismatches 255;	Indels 46;	Gaps 5

Qy 66 GCGGAGTGGCGGACGGGTGAGTAATGATCTGGAGCGTATCCGAAAGAGGGGGGTAAACGA 125

Db 20 GGGGAGTGGCGACCGGGTGAAGTAACCGCTGGGAATCTACCCATCTCTGGGAATAGCTT 79

Qy 126 TCGAAGAATGTGCTAATAACCCGATTAATCTTAAGAGGAAAGCAGGGGATCGAAGACCT 185

Db 80 GGGAAATCTGGAATTAATACCGCATACGCCCTACGGGG-----GAAAGATT 125

QY	186	TGCGTTTGGAGCGCGCGAATGTCGATTACTAGTTGTTGGGGTAAAGGCTTACCAAG	245
Db	126	ATCGGGATGATGAGCCCGGTTGGAATTACTAGTTGTTGGGGTAAAGGCTTACCAAG	185
QY	246	CGACGATCAGTAGTGTCTGAGAGAGAGACACGACGACACTGAGGACTGACACGCGCCA	305
Db	186	CGACGATCAGTAGTGTCTGAGAGAGATGATCAGGACAAATTGGACTGAGACACGCGCCA	245
QY	306	GACTCTTACGGGAGGACAGCACTGAGGAAATTTTGAACAATGGGCGCAAGCCTGATCAGCA	365
Db	246	AACTCTTACGGGAGGACAGCACTGAGGGAATATTGGACAATGGGCGCAAGCCTGATCAGCC	305
QY	366	ATGCCGCTGAGTGAAGAAAGCCTTCGGGTTGTAAAGCTCTTTCACTGAGTGAAGAAAGAAAG	425
Db	306	ATGCCGCTGAGTGAAGAAAGCCTTAAAGGTTGTAAAGCTCTTTCAACGATGAAGA-----	360
QY	426	TTAGCGTAAATATGTAAGCTCATGACGGTATCGACAGAAAGAGACCGGCTTAATACGT	485
Db	361	-----TAAATGACGGTATGTCGAGAAAGAAAGCCCGGCTTAATCTTGT	400
QY	486	GCCAGACAGCCCGGTAATACGTAGGGTGCAAGCGTTAATCGGAAATTACGAGCGTAAAG	545
Db	401	GCCAGACAGCCCGGTAATACGAAAGGCGGCTAGCGTTGTTGGGAATTAATCGAGCGTAAAGC	460
QY	546	GTCGCGACGCGGCTTTGTAAGTCAAGTGTAAATCCCGGCTTAATCTGGGAATTGCGT	605
Db	461	GCACGTAGGCGGATTTTAAAGTCAAGGGGTGAATCCCGCACTCAACCTGCCGAATCGCT	520
QY	606	TTGAACCTACAAAGGCTAGTGTGGGACAGAGGAGGTGGAAATTCATGTTGTAGCAGTGAA	665
Db	521	TTGATACTGGGTATCTTGAATGAGAAAGGTAAAGTGAATTCGAGTGTAGAGGTGAA	580
QY	666	TGCGTAGAGATATGGAAGAACATGATGGCGCAAGGCAAGCCTCTGCGTTAAACATGACGC	725
Db	581	TTTCGTAGATATTCCGAGGAGACACAGTGGGGAAGGCGGCTTAAGTGTCAATTACGACGC	640
QY	726	TCATGACAGAAAGCGTGGGGAGCAACAGATTGATACCTGTGATGTCACGCGCTTAA	785
Db	641	TGAGGTGCGAAAGCGTGGGGAGCAACAGATTGATACCTGTGATGTCACGCGCTTAA	700
QY	786	CGATGTCAACTAGTTGTTGGGCTTTATTAAGCTTGTGTACGAGGTAAACGAGTGAAGTTG	845
Db	701	CGATGAATGTTAGCCGTCGCGGACGATATCTTTGCGTGGCCGAGTTAAACGATTAAACAT	760
QY	846	ACCCTCTGGGAGTACGGTCGCAAGATTAAACTCAAGAAATTGACGAGGACCCGACACA	905
Db	761	TCCGCTTGGGAGTACGGTCCGCAAGATTAAACTCAAGAAATTGACGAGGACCCGACACA	820
QY	906	AGCGGTGATTATGTGATTAAATTCGATGCAACGCGAAAACTTACCTACCTTTGACAT	965
Db	821	AGCGGTGAGCATGTGGTTTAATTTCGAAGCAACCGCGCAACCTTACCGACTTTGACAT	880
QY	966	-----GTAGCGAATTTTCTAAGATTAAGATTAGTCTTCGGGAAGCGTAAACAAGGTGCTG	1020
Db	881	TGCGGGTATGGGCATTTGGAGACGATGTCCTTCAGTTAGGCTGCGCCCAAGAACAGGTGCTG	940
QY	1021	CATGGCTGTGCTCAGCTCGTGTCTGTGAGATTTGGGTTAAAGTCCCGCAACGAGCGCAAC	1080
Db	941	CATGGCTGTGCTCAGCTCGTGTCTGTGAGATTTGGGTTAAAGTCCCGCAACGAGCGCAAC	1000
QY	1081	CTTGTCAATTAAATGCGCATCATTTGGTGGGCACTTTAATGAGACTCCGCGTGAACAACCG	1140
Db	1001	CTCGCCCTTAAGTTCGACGACATTTATGTTGGGCACTTAAGGGGACCTGCGCGGATTAAGCGG	1060
QY	1141	-GAGGAAAGTGGGATGACGTCGAAGTCTCTATGGCCCTTAATGGGTAAGGCTTCAACGTA	1199
Db	1061	AGAGGAAAGTGGGATGACGTCGAAGTCTCTATGGCCCTTAATGGGCTGGGCTACACACGTG	1120
QY	1200	ATACCAATGGCGCTAACAGAGGTTGGCCAAACCGGAGAGGGAGGCTTAATCTCAGAAAGGCG	1259
Db	1121	CTTCAATATGTTGTATCAAGTGGGACGCAAGACAGAGATGTGAGACTTAATCTCCAAAAGC-C	1179
QY	1260	GTCTGATGCTCGGATCGGAGTCTGCAACTCGACTCCGTGAAGTGGGAATTCGCTAGTAATCG	1319

Db 1180 ATCTCAGTTCCGATTTGCACTCTGCAACTCGAGTGCATGAAGTTGAAATGCTAGTAATCG 1239  
Qy 1320 CGGATCAGCATGTCGCGGTAATACGTTCCCGGATCTTGTACACACCGCCCTCAACCA 1379  
Db 1240 CAGATCAGCATGTCGCGGTAATACGTTCCCGGATCTTGTACACACCGCCCTCAACCA 1299  
Qy 1380 TGGAGTGGGTTTCAACGAAAGCAGTACTTAACCGTAAGAGGCGCTTCCACGGTG 1439  
Db 1300 TGGAGTGGGTTTCAACCGAAGTAGTGGCTTAACCGCAAGAGGCGAGCTTAACAGGTA 1359  
Qy 1440 AGATTCAATGACTGGGCTG 1457  
Db 1360 GGGTCAGCGACTGGGCTG 1377

RESULT 54  
US-10-954-147-15  
; Sequence 15, Application US/10954147  
; Publication No. US20050289672A1  
; GENERAL INFORMATION:  
; APPLICANT: CAMBIA  
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells  
; FILE REFERENCE: CAMBIA 414  
; CURRENT APPLICATION NUMBER: US/10/954,147  
; PRIOR FILING DATE: 2004-09-28  
; PRIOR APPLICATION NUMBER: 60/583,426  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 15  
; LENGTH: 1433  
; TYPE: DNA  
; ORGANISM: A. tumefaciens EHA105  
US-10-954-147-15

Query Match 57.1%; Score 832; DB 10; Length 1433;  
Best local Similarity 77.8%; Pred. No. 9,4e-250;  
Matches 1087; Conservative 0; Mismatches 265; Indels 46; Gaps 5;

Qy 66 GCGCAGTGGCGGACGGGTGAGTAATGCATCGGAACGATTCGAAGAAGGGGGGTAACCGCA 125  
Db 20 GGGGAGTGGCGAGACGGGTGAGTAACGGTGGGAATCTACCCATCTCTGCGAAATAGCTCT 79  
Qy 126 TCGAAGATGTGCTAATACCGCATATATCTTAAGAGGAAGCAGGGGATCGAAAGCT 185  
Db 80 GGGAACTGGAATTAATACCGCATACGCCCTTACGGG-----GAAAGATT 125  
Qy 186 TGGCTTTTGGAGCGGCCGATGCTGATTAAGTATGTTGGGTAAAGGCTTAACCAAG 245  
Db 126 ATGGGGATGAGTAGAGCCCGCTTGGATTAAGTATGTTGGGTAAAGGCTTAACCAAG 185  
Qy 246 CGACGATCAGTATGTTGTTGAAGAGCAGCAGCCACATCGGAGCTGAACACGGCCCA 305  
Db 186 CGACGATCAGTATGTTGTTGAAGAGTATCAAGCCCATTTGGACTGAACACGGCCCA 245  
Qy 306 GATCTCTACGGAGGAGCAGTGGGAAATTTGGACAATGGGGCAAGCCGATTCAGCA 365  
Db 246 AATCTCTACGGAGGAGCAGTGGGAAATTTGGACAATGGGGCAAGCCGATTCAGCA 305  
Qy 366 ATGCGCGGTAGTGAAGAAGGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAAGAAAG 425  
Db 306 ATGCGCGGTAGTGAAGAAGGCTTCAGGTTGTAAAGCTCTTTCACGAGTAAGAA----- 360  
Qy 426 TTAAGTAAATTAATCGTACTCATGACGGTATCGACAGAAAGACCCGGCTTAATCTGCT 485  
Db 361 -----TAATGACGGTATGTCGAGAAAGAACCCCGGCTTAATCTGCT 400  
Qy 486 GCCAGCAGCCGCGGTTAATAGTAAAGGTCGAAAGCTTAATCGGATTTACTGGGCGTAAG 545  
Db 401 GCCAGCAGCCGCGGTTAATAGTAAAGGTCGAAAGCTTAATCGGATTTACTGGGCGTAAG 460  
Qy 546 GTGCGCAGCGCGGCTTTTAAGTCAAGATGTGAATCCCGGCGCTTAACCTGGAAATGCGT 605

Db 461 GCAGTAGGCGGATATTAATGTCAGGGGTGAAATCCCGAGCTCAACTGCCGAATGCTGCT 520  
Qy 606 TTGAAATCAAGAGCTAGAGTGTGGCAGAGGAGTGGAAATTCATGTTGATGAGTAA 665  
Db 521 TTGATACTGGGTATCTTGATGATGAAGAGGTAAAGTGAATTCGAGTTGAGAGTAA 580  
Qy 666 TGGCGATGATGATGAAGAAATCATGATGCGAAGGCAAGCCCTCTGGGTTAACTAGACG 725  
Db 581 TTGATGATATTCGAGGAACACAGTGGCGAAGGCGGCTTACTGCTCATTAATGACGC 640  
Qy 726 TCATGCAAGAAAGCGTGGGAGCAAAAGATTAAGATACCTTGGTATGTCACGCCCTTAA 785  
Db 641 TGAAGTGAAGAAAGCGTGGGAGCAAAAGATTAAGATACCTTGGTATGTCACGCCCTTAA 700  
Qy 786 CGATGCAATGATGTTGTTGGGCTTATTAAGCTTGTATGAAGAGCTAAGGCTGAAGTTG 845  
Db 701 CGATGATGTTAGCCGTGGGCGATATCTGTTGGTGGCGAGCTTAACGATTAACAT 760  
Qy 846 ACCGCTTGGGAGTACGGTTCGCAAGATTAATACTCAAGAAATTAAGCGGAGCCGCA 905  
Db 761 TCCGCTGGGAGATACGGTTCGCAAGATTAATACTCAAGAAATTAAGCGGAGCCGCA 820  
Qy 906 AGCGGTGATTAATGATTAATTCGATGCAACGCAAAACCTTAACCTTACCTGACAT 965  
Db 821 ACCGTTGAGCATGTGCTTTAATGAGCAACGCGCAAACTTAACGATCTTGAACAT 880  
Qy 966 -----GTAGCGAATTTTCTAGATATGATTAAGTCTTGGGAAACGCTAACACAGTCTG 1020  
Db 881 TCGGGGTATGGGCATTTGAGACGATGCTCTTCAGTTAGGCTGGCCCAAGAGTCTG 940  
Qy 1021 CATGCTGTGCTACCTGCTGTCGATGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1080  
Db 941 CATGCTGTGCTACCTGCTGTCGATGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1000  
Qy 1081 CTTGTCAATTAATGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Db 1001 CTCGCCCTTAATGTCAGCATTTAGTTGGGCACTTAAGGGGACGCGGTATTAAGCCG 1060  
Qy 1141 -GAGGAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199  
Db 1061 AGAGGAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1120  
Qy 1200 ATACATGCGCGCTACAGAGGTTGCCAACCGCGAGGGGAGCTTAATCTCAAGAAAGCCG 1259  
Db 1121 CTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1179  
Qy 1260 GTGATGTCGAGTGGAGTCTGCAACTGCACTCCGTGAAGTCGAAATGCTAATATCG 1319  
Db 1180 ATCTCAGTTCCGATTTGCACTCTGCAACTGCACTGCAATGAAGTTGGAATGCTAATATCG 1239  
Qy 1320 CGGATCAGCATGTCGCGGTAATACGTTCCCGGATCTTGTACACACCGCCCTCAACCA 1379  
Db 1240 CAGATCAGCATGTCGCGGTAATACGTTCCCGGATCTTGTACACACCGCCCTCAACCA 1299  
Qy 1380 TGGAGTGGGTTTCAACGAAAGCAGTACTTAACCGTAAGAGGCGCTTCCACGGTG 1439  
Db 1300 TGGAGTGGGTTTCAACCGAAGTAGTGGCTTAACCGCAAGAGGCGAGCTTAACAGGTA 1359  
Qy 1440 AGATTCAATGACTGGGCTG 1457  
Db 1360 GGGTCAGCGACTGGGCTG 1377

RESULT 55  
US-11-055-637-84  
; Sequence 84, Application US/11055637  
; Publication No. US20050289619A1  
; GENERAL INFORMATION:  
; APPLICANT: BROUSSEAU, Roland  
; APPLICANT: DUBOIS, Jason  
; APPLICANT: EDGE, Tom  
; APPLICANT: MASSON, Luc

```

; APPLICANT: TREVORS, Jack T.
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
; TITLE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES
; FILE REFERENCE: 2139-3305
; CURRENT APPLICATION NUMBER: US/11/055,637
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,288
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Probe for DNA array
; US-11-055-637-84

Query Match      57.1%; Score 831.8; DB 17; Length 1470;
Best Local Similarity 76.9%; Pred. No. 1.1e-249;
Matches 1080; Conservative 2; Mismatches 309; Indels 14; Gaps 5;

Qy 66 GCGAGTGGCGGACGGGTGAGTATGATCGG-AAAGTATCCAGAAAGGGGGGTAACGC 124
Db 55 GGTAGCGCGGACGGGTGAGTATGATCGGAACTGCGCTTAAGACGCGGATTAATCTC 114
Qy 125 ATCGAAATATGCTTAATACCGCATATCTCTAAGAGAAAGCAGGGGATCGAAG--- 181
Db 115 ACGGAAGGTGGGTAAAGACGGATGAGCGATTTGCTCGCATGAGGAAATCGGGAAGGC 174
Qy 182 -----ACCTTGCGCTTTTGGAGCGCGGATGTCTGATTTAGTTAGTGGGGTAAAG 235
Db 175 GGAGAAATCTGCGCACTTATGATGATGACCTACGCGCCATTAGTATGATGATGATGATG 234
Qy 236 CCTACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 295
Db 235 CTCACCAAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 294
Qy 296 ACACGGCCCAAGCTCTTACGGGAGGACAGTGGGGAATTTTGGACATTTGGCGGCAAGCC 355
Db 295 ACACGGCCCAAGCTCTTACGGGAGGACAGTGGGGAATTTTGGACATTTGGCGGCAAGTC 354
Qy 356 TGATCCAGCAATGCGCGGAGTGAAGAGGCTTCGGGGTTTAAAGCTTTTCAGTGA 415
Db 355 TGAACGAGCAAGCGCGGAGTGAAGAGGCTTCGGGGTTTAAAGCTTTTCAGTGA 414
Qy 416 GAAGAAAGGTTACCGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 475
Db 415 GAAGAAAGGTTACCGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 474
Qy 476 CTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 535
Db 475 CTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 534
Qy 536 GACGTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 595
Db 535 GACGTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 593
Qy 596 GGAATTGGCTTGAATCTACAGGCTAGAGTGTGACAGAGGAGGTGATTCATGTGT 655
Db 594 CGGGTCGATCGAAATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 653
Qy 656 AGCAGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 715
Db 654 AGCAGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 713
Qy 716 ACATGACGCTCATGACAGAAAGGTTGGGAGCAAAAGATTTAGATACCCCGGTAGTCC 775
Db 714 TAACTGAGCTGAGGCGGAAAGGTTGGGAGCAAAAGATTTAGATACCCCGGTAGTCC 773
Qy 776 ACGCCCTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 833
Db 774 ACGCCCTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 833
```

```

Qy 834 GCGGGAAGTGTACCGCTGGGAGTACCGTCCGAAGATTAAACTCAAGAAATTGACG 893
Db 834 CACTTAAAGCATTTCCGCTGGGAGTACCGTCCGAAGATTAAACTCAAGAAATTGACG 893
Qy 894 GGAACCGGACCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 953
Db 894 GGAACCGGACCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 953
Qy 954 TACCTTGACATGTAAC-GAATTTTCTAGATGATGATGATGATGATGATGATGATGATGATG 1012
Db 954 AGGTCTTGACATGTAACCGCTGACCGGCTAGAGATGAGGCTTCCCTTGGGCGAGAGAGAC 1013
Qy 1013 AGGTCTTGACATGTAACCGCTGACCGGCTAGAGATGAGGCTTCCCTTGGGCGAGAGAGAC 1072
Db 1014 AGGTCTTGACATGTAACCGCTGACCGGCTAGAGATGAGGCTTCCCTTGGGCGAGAGAGAC 1073
Qy 1073 GCGCAACCTTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1132
Db 1074 GCGCAACCTTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1133
Qy 1133 ACAAACCGGAGAAAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1192
Db 1134 ACAAACCGGAGAAAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1193
Qy 1193 ACAGTATATCAATGAGGCGGTACAGAGGCTTCCCAACCGGAGGAGGAGCTTAATCTCAG 1252
Db 1194 ACAGTATATCAATGAGGCGGTACAGAGGCTTCCCAACCGGAGGAGGAGCTTAATCTCAG 1253
Qy 1253 AAACGCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1312
Db 1254 AAACGCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1313
Qy 1313 GTAAATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1372
Db 1314 GTAAATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1373
Qy 1373 CACACATGAGGAGGAGGTTTACCAAGAGAGGATGATGATGATGATGATGATGATGATGATGATG 1432
Db 1374 CACACATGAGGAGGAGGTTTACCAAGAGAGGATGATGATGATGATGATGATGATGATGATGATG 1433
Qy 1433 CACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1457
Db 1434 GAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1458

RESULT 56
US-10-953-392-10
; Sequence 10, Application US/10953392
; Publication No. US20050289667A1
; GENERAL INFORMATION:
; APPLICANT: CAMBIA
; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells
; FILE REFERENCE: CAMBIA 414B
; CURRENT APPLICATION NUMBER: US/10/953,392
; CURRENT FILING DATE: 2004-09-28
; PRIOR APPLICATION NUMBER: US 60/583,426
; PRIOR FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 1368
; TYPE: DNA
; ORGANISM: P. myrsinacearum Cambia isolate
US-10-953-392-10

Query Match      57.0%; Score 830.8; DB 10; Length 1368;
Best Local Similarity 78.7%; Pred. No. 2.2e-249;
Matches 1085; Conservative 0; Mismatches 247; Indels 46; Gaps 6;

Qy 66 GCGAGTGGCGGACGGGTGAGTATGATCGGAACTGATCCAGAAAGGGGGGTAACGCA 125
Db 31 GCGAGTGGCGGACGGGTGAGTATGATCGGAACTGATCCAGAAATGTAACGCA 90
```

QY 126 TCGAAGATGTGCTAATACCCGATATCTTAAGAGGAAAGCAGGGGATCGAAAGCCT 185  
 DB 91 TGGAAAGCTGTGCTAATACCGATATCGTCTCGGGA-----GAAAGATT 136  
 QY 186 TGGCTTTTGAAGCCGCGATCTGATTAGTCTGTTGGGTAAAGGCTTACCAAG 245  
 DB 137 ATCGAAGATGATGAGCCCGCGTGGATTAGCTAGTGGGTAAAGGCTTACCAAG 196  
 QY 246 GCAGCATCATGATGTGCTGAGAGAGCAGCAGCACTGGGACTGAGACACGGCCA 305  
 DB 197 GCAGCATCATGATGTGCTGAGAGATGATCAGCAGCTGGGACTGAGACACGGCCA 256  
 QY 306 GACTCTTACGAGGAGCAGCAGTGGGAAATTTTGAACAATGGCGCAGAGCTGATCAGCA 365  
 DB 257 GACTCTTACGAGGAGCAGCAGTGGGAAATTTGACAATGGCGCAGAGCTGATCAGCC 316  
 QY 366 ATGCGCGGTGATGAGAAAGCCTTCGGGTTTAAAGCTCTTCACTGAGAGAAAG 425  
 DB 317 ATGCGCGGTGATGAGAAAGCCTTACGGGTTTAAAGCTCTTCACTGAGAGAAAG----- 371  
 QY 426 TTAAGCTAATATCTGATCTGATGAGCGGATCGACAGAAAGACACCGGCTTACCTAGT 485  
 DB 372 -----TATGACGGTAAACCGGAGAGAAAGCCCGGCTTACCTAGT 411  
 QY 486 GCAGAGCGCGGCTAATACGTAGGCTGACAGCGTTAATCGGAATTACTGGCGTAAAG 545  
 DB 412 GCCAGAGCGCGGCTAATACGAAGGGGCTAGGTTGTTCCGATTACTGGCGTAAAGC 471  
 QY 546 GTGCGAGCGCGCTTGTAAAGTCAATGTGAAATCCCGGCTTAACTGGGAAATGGCT 605  
 DB 472 GCACGTAGCGGATTAAGTCAAGGGGTGAATCCCGGCTTAACTGGGAAATGGCT 531  
 QY 606 TTGAAATCTAAGGCTAAGTGTGAGAGGAGGAGGAAATTCAGTGTAGAGAGAA 665  
 DB 532 TTGATCTGTAGTCTTGTAGTTCGAGAGGAGTGAATTCAGTGTAGAGAGAA 591  
 QY 666 TGGGTGAGATTAAGAGAAATCATGATGCGGAGAGGAGCTCTGGGTTAACTAGTACGC 725  
 DB 592 TTGATGATTAATTCGAGAGAAATCATGATGCGGAGAGGAGCTCTGGGTTAACTAGTACGC 651  
 QY 726 TCAATGACGAAAGCGTGGGAGCAAAAGAGATTAGTATCCGTGATGCTCAAGCCTTAA 785  
 DB 652 TGAAGTGAAGACGTTGGGAGCAAAAGAGATTAGTATCCGTGATGCTCAAGCCTTAA 711  
 QY 786 CGATGCTAATGATGTTGGGCTTATAGGCTTGTGTAAGCAAGCTTAAAGCGGAGATTG 845  
 DB 712 CTATGAGAGCTACCGCTCGGAGATTAATGTTGCGGTGGGAGCAAAAGCATTAAAGCTC 771  
 QY 846 ACCGCTGGGAGATAGCGTGCAGAAATTAAACTCAAGGAATTGACGGGAGACCGGACA 905  
 DB 772 TCGGCTGGGAGATAGCGTGCAGAAATTAAACTCAAGGAATTGACGGGAGACCGGACA 831  
 QY 906 AGCGGTGATTAATGATTAATTCATGAGCAAGCGGAAACCTTACCTTACCTTGCAT 965  
 DB 832 AGCGGTGATTAATGATTAATTCATGAGCAAGCGGAAACCTTACCTTGCAT 891  
 QY 966 --GTAGGAAATTTCTAGAGATGATTAATG--GCTTCGGGAAAGCTTAAAGCAAGTGTG 1020  
 DB 892 CCCGATCGCGGATTAACAGAGATGTTTCTTCAATTAAGGCTGATTCGTTGAGAGAGTGTG 951  
 QY 1021 CATGGCTGTGCTGAGCTGCTGAGATGTTGGTTAAGTCCCGGACGAGCGGACAC 1080  
 DB 952 CATGGCTGTGCTGAGCTGCTGAGATGTTGGTTAAGTCCCGGACGAGCGGACAC 1011  
 QY 1081 CTGTGCTAATTAATGCTAATTTGGTGGGCACTTAAATGAGACTGCGGTGACAAACG 1140  
 DB 1012 CTGCGCTTAAATGCTAATTTGGTGGGCACTTAAATGAGACTGCGGTGACAAACG 1071  
 QY 1141 -GAGGAAAGTGGGATGAGCTCAAGTCTCAATGGGCTTAAATGAGGATGAGGCTTCAAGCTA 1199  
 DB 1072 AGAGGAAAGTGGGATGAGCTCAAGTCTCAATGGGCTTAAATGAGGATGAGGCTTCAAGCTA 1131

QY 1200 ATACATGCGCGGATACAGAGGTTGCCAACCCCGGAGGGGAGCTAATCTCAAGAACGC 1259  
 DB 1132 CTACATATGCTGTGATCAGTGGGACGACGACCGGAGGTGAGCTAATCTCAAGAACGC 1190  
 QY 1260 GTCTAGTCCGGATCGAGTCTGCACTCGATCTCGTGAAGTGGAAATCGCTAATATCG 1319  
 DB 1191 ATCTCAGTGGGATTTGCACTCTGCACTCGAGTGCATGAAATGGAATCGCTAATATCG 1250  
 QY 1320 CGATACAGATGTGCGGCTGAATACGTTCCCGGCTCTTGAACACACCGCCGTCACACA 1379  
 DB 1251 TGGATTCAGAAATGCCAGGTGAATACGTTCCCGGCTCTTGAACACACCGCCGTCACACA 1310  
 QY 1380 TGGAGTGGGTTTACAGGAGAGAGTATTAACCTGATAGAGAGCGCTTGCACCG 1437  
 DB 1311 TGGAGTGGTTTTACCGAAGGTGCTGTCTTAACCGAAGAGAGGCAACACCG 1368

RESULT 57  
 US-10-954-147-10  
 ; Sequence 10, Application US/10954147  
 ; Publication No. US20050289672A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CAMBIA  
 ; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells  
 ; FILE REFERENCE: CAMBIA 414  
 ; CURRENT APPLICATION NUMBER: US/10/954,147  
 ; CURRENT FILING DATE: 2004-09-28  
 ; PRIOR APPLICATION NUMBER: 60/583,426  
 ; PRIOR FILING DATE: 2004-06-28  
 ; NUMBER OF SEQ ID NOS: 83  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 10  
 ; LENGTH: 1368  
 ; TYPE: DNA  
 ; ORGANISM: P. myrsinacearum Cambia isolate  
 US-10-954-147-10

Query Match 57.0%; Score 830.8; DB 10; Length 1368;  
 Best Local Similarity 78.7%; Pred. No. 2.2e-249;  
 Matches 1085; Conservative 0; Mismatches 247; Indels 46; Gaps 6;

QY 66 GCGAGTGGCGGACGGGTGATGATTCGATCGGAGATTCGAAAGGGGATTAACGCA 125  
 DB 31 GGGAGCGGACAGCGGTGATGATTCGATCGGAGATTCGAAAGGGGATTAACGCA 90  
 QY 126 TCGAAGATGTGCTAATACCGCATTAATCTTAAGAGGAAAGCAGGGGATCGAAAGCCT 185  
 DB 91 TGGAAAGCTGTGCTAATACCGATATCGTCTCGGGA-----GAAAGATT 136  
 QY 186 TGGCTTTTGAAGCCGCGATCTGATTAGTCTGTTGGGTAAAGGCTTACCAAG 245  
 DB 137 ATCGAAGATGATGAGCCCGCGTGGATTAGCTAGTGGGTAAAGGCTTACCAAG 196  
 QY 246 GCAGCATCATGATGTGCTGAGAGAGCAGCAGCACTGGGACTGAGACACGGCCA 305  
 DB 197 GCAGCATCATGATGTGCTGAGAGATGATCAGCAGCTGGGACTGAGACACGGCCA 256  
 QY 306 GACTCTTACGAGGAGCAGCAGTGGGAAATTTTGAACAATGGCGCAGAGCTGATCAGCA 365  
 DB 257 GACTCTTACGAGGAGCAGCAGTGGGAAATTTGACAATGGCGCAGAGCTGATCAGCC 316  
 QY 366 ATGCGCGGTGATGAGAAAGCCTTCGGGTTTAAAGCTCTTCACTGAGAGAAAG 425  
 DB 317 ATGCGCGGTGATGAGAAAGCCTTACGGGTTTAAAGCTCTTCACTGAGAGAAAG----- 371  
 QY 426 TTAAGCTAATATCTGATCTGATGAGCGGATCGACAGAAAGACACCGGCTTACCTAGT 485  
 DB 372 -----TATGACGGTAAACCGGAGAGAAAGCCCGGCTTACCTAGT 411  
 QY 486 GCAGAGCGCGGCTAATACGTAGGCTGACAGCGTTAATCGGAATTACTGGCGTAAAG 545  
 DB 412 GCCAGAGCGCGGCTAATACGAAGGGGCTAGGTTGTTCCGATTACTGGCGTAAAGC 471



```
Qy 546 GTGCCGAGGCGGCTTGTAGTCAAGTCAATGTAATCCCGGGCTTAACCTGGAAATTGCGT 605
Db 472 GCACGTAGCGGCACTATTAAAGTCAAGGGGTGAATCCCGGGCTCAACCCCGGAACCTGCT 531
Qy 606 TTGAAACTACAAAGGCTAGAGTGTGCGAAGGGAAGTGGAAATTCATGTGTACAGTGA 665
Db 532 TTGATATCTGTAGTCTTGTGAGTTGAGAGAGGTGAAGTAATCCGAGTGTAGAGGTGA 591
Qy 666 TGGCGAGATATGGAAGAACATCGATGCGAAGGAGCGCTCTGGGTTAACAATGACGCG 725
Db 592 TTCTGATGATATTCGAGAGAACACAGTGGCGAAGGGGCTCACTGCTCGATATCTGACGC 651
Qy 726 TCATCAGCAAGAACGCTGGGAGCAAAACAGATTAAGTATCCTGTAGTCCAGCCCTAA 785
Db 652 TGAAGTGCAGAAAGCGTGGGAGCAAAACAGATTAAGTATCCTGTAGTCCAGCCCTAA 711
Qy 786 CGATGTCAAATGATGTTGGGCTTATTAAGGCTTGTGTAACGAGCTAACGCGTGAAGTTG 845
Db 712 CTATGAGAGCTAGCGGTGGGAGATATCTGTTCCGTGGCGCAGCAAAACGATTAAGCTC 771
Qy 846 ACCGCTGGGAGTACGCTGCGAAGATTAAACTCAAGGAATTGACGGGGACCCGCA 905
Db 772 TCCGCTGGGAGTACGCTGCGAAGATTAAACTCAAGGAATTGACGGGGACCCGCA 831
Qy 906 AGCGGTGATATGATGATTAATTCGATGCAACGCAAAACCTTAACCTACCTGACAT 965
Db 832 AGCGGTGATGATGATGATTAATTCGATGCAACGCGCAAAACCTTAACCTGACAT 891
Qy 966 --GTAGCGAATTTTGAAGATGATTAAGT--GCTTGGGAAACGCTAACACAGGTGCTG 1020
Db 892 CCCGATCCGCGTTACAGAGATGATGATTTCTTCAATTAAGCTGATGCGGTGAGTGC 951
Qy 1021 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 952 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1011
Qy 1081 CTGTCATTAATTCGATCAATTTGGTGGGCACTTAAATGAGATGCGCGGTGACAAACG 1140
Db 1012 CTCGCCCTTATGTCGATCAATTTGGGCACTTAAATGAGATGCGCGGTGACAAACG 1071
Qy 1141 -GAGGAGGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGAGGTGAGGGCTTACACGTA 1199
Db 1072 AGAGGAGGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGAGGTGAGGGCTTACACG 1131
Qy 1200 ATACATGCGCGGTGACAGAGGTGTCCAACCCGCGAGGGGAGCTTAATCTCAAGAAAGCG 1259
Db 1132 CTACATGCGGTGAGTGAAGTGGGACGAGACCGCGAGGTGAGCTTAATCTCAAGAAAGCG 1190
Qy 1260 GTGCTAGTCCGAGTGGAGTCTGCAACTGCACTCCGTGAAGTGGAGTGGATGATG 1319
Db 1191 ATCTCAATGTCGATGTCGCACTGCAACTGCACTGCAATGTCGATGTCGATGTCG 1250
Qy 1320 CGGATCAGCATGTCGCGGTGAATACGTTCCCGGCTTGTGACACACGCGCGTCAACCA 1379
Db 1251 TGGATCAGAAATGCGCAGGTAATCGTTCCCGGCTTGTGACACACGCGCGTCAACCA 1310
Qy 1380 TGGAGTGGGTTTCAACGAGACAGTGTCTAACCGTGAAGAGGGCGCTTGCACCG 1437
Db 1311 TGGGAGTTGTGTTTACCGGAAGTGTCTGTCTAACCGGAGAGGACAGCAACCG 1368

RESULT 58
US-11-055-637-83
; Sequence 83, Application US/11055637
; Publication No. US20050260619A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: DUBOIS, Jason
; APPLICANT: EDER, Tom
; APPLICANT: MASSON, Luc
; APPLICANT: TREVORS, Jack T.
; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND
; TITLE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES
```

```
/ FILE REFERENCE: 2139-33US
/ CURRENT APPLICATION NUMBER: US/11/055.637
/ PRIOR FILING DATE: 2005-02-11
/ PRIOR APPLICATION NUMBER: US 60/543,288
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 83
/ LENGTH: 1491
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Probe for DNA array
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1491)
/ OTHER INFORMATION: n = A,T,C or G
US-11-055-637-83

Query Match 56.9%; Score 828.6; DB 17; Length 1491;
Best Local Similarity 75.5%; Pred. No. 1.1e-248;
Matches 1114; Conservative 0; Mismatches 340; Indels 22; Gaps 6;

Qy 4 GAAGCTGGCGGCAATGCTTTACATGCAAGTCGAACGGGACAGCAGATGCTTCATCTG 63
Db 4 GAAGCTGGCGGCGGCTTCATATACATGCAAGTCGAACGGGACAGCAGATCTCTT 63
Qy 64 GTGCGAGTGG-----CGACGGGTGAGTATATGATCGG-AACGATCCGAAGAG 113
Db 64 CGGGGTGACCTGTTGTCMAACGGGTGAGTAAACGTGGGCAACCTTACTGTAAAGC 123
Qy 114 GGGGTAAACGATGAAGATGTCATTAATCCGATATACCTTAAGAGGAAACAGGGG 173
Db 124 TGGATTAATCTCGGAAACCGGGCTTAATACCGATATACATATCTTCCATACAGATG 183
Qy 174 ATCGAAAGAC-----CTTGCGCTTTTGAAGCGCGCGATGTCGATTAAGTGGT 225
Db 184 TTGAAAGGGCGGCAATGCTGCTCACTTAACAGATGGCCCGGCGCATTAAGTGGT 243
Qy 226 GGGTAAAGGCTTACAAAGGATGATGATGATGATGATGATGATGATGATGATGATG 285
Db 244 GAGATAAAGGCTTACAAAGGATGATGATGATGATGATGATGATGATGATGATGATG 303
Qy 286 TGGGATGAGACAGGGCCAGATCTTACGGGAGGACGATGAGGGAATTTTGGACATG 345
Db 304 TGGGATGAGACAGGGCCAGATCTTACGGGAGGACGATGAGGGAATTTTGGACATG 363
Qy 346 GGCAGAGCTGATCCAGCAATGCGCGTGAAGTGAAGAGGCTTCCGTTGTAAGCTC 405
Db 364 GAGCAAGTCTGACGAGCAACCGCGGTGATGATGAAGTTTTCGATCGTAAATCTC 423
Qy 406 TTTCAATGAGAGAAAGGTTACGATTAATATGATGATCA-TGACGATATGACAGA 464
Db 424 TGTGTGAGGAGAAAGAAAGTGCATTCGATTAATGTTGGACCTTGAACGCTTAACCA 483
Qy 465 AGAAGACCGGCTTAACGATGCGCAGACCGGGTAAATGATGAGGTGAAGGCTTAAT 524
Db 484 GAAAGCCCGCTTAACGATGCGCAGACCGGGTAAATGATGAGGTGAAGGCTTATC 543
Qy 525 CGGAATTAATGAGGCTTAAAGGAGTGGCGGCTTGTGAAGTCAATGTAATCCCG 584
Db 544 CGGAATTAATGAGGCTTAAAGGAGTGGCGGCTTGTGAAGTCAATGTAATCCCG 603
Qy 585 GGCCTTAACCTGGGAATTCGTTGAAACTCAAGGCTTAAGTGTGCGAGAGGAGTGA 644
Db 604 GGCCTTAACCTGGAGGAGCTTGAAGCTGGGGAATTAAGTCAAGAGAGGAGTGA 663
Qy 645 ATTCAATGATGAGAGGAAATGCTGATGATGAGAAACATCAATGCGGAAGGAGC 704
Db 664 ATTCAATGATGAGAGGAAATGCTGATGATGAGAAACATCAATGCGGAAGGAGC 723
Qy 705 CTCCTGGTTAACACTACGCTCATGCAAGAAAGCGGGAGCAACAGATTAAGTAC 764
```



Db 724 TCTGTGCTGTAACTGACGCTGAGTCCGAAAGCGTGGGTAAGCAACAGATTAGATAC 783  
 Qy 765 CCGTGTAGTCCACGCGCTTAAAGATGTCACTAGTTGTT--GGGCTTATTAGGCTTGGT 822  
 Db 784 CCGTGTAGTCCACGCGCTTAAAGATGTCACTAGTTGTT--GGGCTTATTAGGCTTGGT 843  
 Qy 823 AACGAGCTTAAAGCGTGAAGTTGACCGCTGGGGAGTACCGTCCGAAATTAAACTCA 882  
 Db 844 GCTGAAGTTAAAGCACTTAAGCACTCCGCTGGGGAGTACCGTCCGAAATTAAACTCA 903  
 Qy 883 AGGAATTGACGGGAGCCCGCAACAGCGGTGATTATGAGATTAAATGATGCAAGCGCA 942  
 Db 904 AGGAATTGACGGGAGCCCGCAACAGCGGTGATTATGAGATTAAATGATGCAAGCGCA 963  
 Qy 943 AAAACCTTACCTTGAACCTTGAACATGTAGCGAATTTTCTAGAGATTAGT-GCTTCCGG 1001  
 Db 964 AGAACCTTACCTTGAACCTTGAACATGTAGCGAATTTTCTAGAGATTAGT-GCTTCCGG 1023  
 Qy 1002 AACGCTAACACAGTGTGCTGATGCTGTGCTGAGCTGCTGTGAGATGTTGGGTTAAG 1061  
 Db 1024 GACACAGTGAACAGTGTGCTGATGCTGTGCTGAGCTGCTGTGAGATGTTGGGTTAAG 1083  
 Qy 1062 TCCGCAACGAGCGCAACCTTGTCTATTATTTGCTCATTTTGGGCTTTAATGA 1121  
 Db 1084 TCCGCAACGAGCGCAACCTTGTCTATTATTTGCTCATTTTGGGCTTTAATGA 1143  
 Qy 1122 GACTGCGGCTGACAAACCGGAGGAGGAGGAGTGAAGTCAAGTCCCTCATGGCCCTTATG 1181  
 Db 1144 GACTGCGGCTGACAAACCGGAGGAGGAGGAGTGAAGTCAAGTCCCTCATGGCCCTTATG 1203  
 Qy 1182 GGTAGGAGCTTCAACGCTAATTAATGAGCGCTGACAGAGGTTGCCAAACCGCGAGGGGA 1241  
 Db 1204 ACTGAGGCTTCAACGCTGCTTCAATGATGAGAAAGGAGCGAGGCGAGGCGCA 1263  
 Qy 1242 GCTAATCTCAAGAAAGCGCTGCTGATGCTGAGTGTGCACTGCACTCCGTTAAT 1301  
 Db 1264 GCAAAATCCCATAAACCATTTCTCAATTCAGATTGCGATGCGGCTGCAACTGCTGATGAAC 1323  
 Qy 1302 CGGAATCGCTAGTAATCGGCGATCAGATGTGCGGCTGATGCTGCGGCTGCTGAT 1361  
 Db 1324 CGGAATCGCTAGTAATCGGCGATCAGATGTGCGGCTGATGCTGCGGCTGCTGAT 1383  
 Qy 1362 ACACGCGCGCTGACACCATGAGGAGTGGTTTCAACGAAACAGATGCTTAACCGTAAG 1421  
 Db 1384 ACACGCGCGCTGACACCATGAGGAGTGGTTTCAACGAAACAGATGCTTAACCGTTTGG 1443  
 Qy 1422 AGGCGCTTGCACCGGTGAGATTGATGCTGGGCTG 1457  
 Db 1444 AGCGAGCGCGCGAAGGTGGGACTAATGATTGGGCTG 1479

RESULT 59  
 US-10-953-392-4  
 ; Sequence 4, Application US/10953392  
 ; Publication No. US20050289667A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CAMBIA  
 ; TITLE OF INVENTION: Biological Gene Transfer System for Eukaryotic Cells  
 ; FILE REFERENCE: CAMBIA 414B  
 ; CURRENT APPLICATION NUMBER: US/10/953,392  
 ; CURRENT FILING DATE: 2004-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/583,426  
 ; PRIOR FILING DATE: 2004-06-28  
 ; SOFTWARE: PatentIn version 3.2  
 ; NUMBER OF SEQ ID NOS: 83  
 ; SEQ ID NO 4  
 ; LENGTH: 1395  
 ; TYPE: DNA  
 ; ORGANISM: S. meliloti 1021  
 ; US-10-953-392-4

Query Match 56.8%; Score 827.2; DB 10; Length 1395;  
 Best Local Similarity 77.5%; Pred. No. 3e-248;

Matches 1084; Conservative 0; Mismatches 268; Indels 46; Gaps 5;  
 Qy 66 GCGAGTGGCGGAGCGGTGAGTATGCAATCGGAACGATCCAGAAAGGAGGAGTAAAGCA 125  
 Db 33 GGGAGCGCGGAGCGGTGAGTAAAGCGTGGAAATCTACCTTTTCTACGGAATTAAGCA 92  
 Qy 126 TCGAAAGTGTCTAATACCGCATATCTTAAGAGGAAAGCAGGGATCGAAAGACT 185  
 Db 93 GGGAAACTTGTCTAATACCGCATATCTTAAGAGGAAAGCAGGGATCGAAAGACT 147  
 Qy 186 TGCCTTTTGGAGCGCGGATGCTGATTAGCTGATTGTTGGTGGGTTAAAGGCTTACCAAG 245  
 Db 148 -----GGATGAGCCCGGCTTGAATTAAGTCTGTTGGTGGGTTAAAGGCTTACCAAG 198  
 Qy 246 GAGCATGATGATTGTTGCTGAGAGACGACGACGACCTGAGGAGCTGAGACGAGGCCA 305  
 Db 199 CGACGATCATATGCTGCTGCTGAGAGGATCAGCAATTGGGACTGAGACGAGGCCA 258  
 Qy 306 GACTCCTTACGGGAGGACGAGTGGGGAATTTTGGACATGAGGCGCAACCTGATCAGCA 365  
 Db 259 AACTCTTACGGGAGGACGAGTGGGGAATTTTGGACATGAGGCGCAACCTGATCAGCC 318  
 Qy 366 ATGCGCGTGAAGTGAAGAGCGCTTGGGTTGAAGCTTTTCACTGAGAGAGAAAGG 425  
 Db 319 ATGCGCGTGAAGTGAAGAGCGCTTGGGTTGAAGCTTTTCACTGAGAGAGAAAGG 373  
 Qy 426 TTACGTAATTAATGCTGATCATGACGATTCAGACAGAAAGACCGGCTTAATCACT 485  
 Db 374 -----TAAATGACGTTAACCGGAGAGAGAGCGGCTTAATCACT 413  
 Qy 486 GCCACAGCGCGGCTTAATCACTGAGGCTCAAGCGTTAATCGGAATTACTGAGCGTAAAG 545  
 Db 414 GCCACAGCGCGGCTTAATCACTGAGGCTCAAGCGTTAATCGGAATTACTGAGCGTAAAG 473  
 Qy 546 GTGCGCAGCGCGCTTGTGAAGTCAAGATGGAATCCCGGCTTAACTGGAAATTCGCT 605  
 Db 474 GCACTGAGCGGATTTGTTAAGTGAAGGAGGATCCAGAGGCTCAACCTGGAACCTGCT 533  
 Qy 606 TTGAATCTAAGGCTTGAAGTGTGCAAGGAGGAGTGAATTCATGCTTACAGTGA 665  
 Db 534 TTCAATGCTGCAATCTGAAGTGCAGAAAGGAGTGAATTCATGCTTACAGTGA 593  
 Qy 666 TCGCTGAGATTAATGAAGAACATGATGCGGAAGGAGGCTCTGGGTTAACTGACGCG 725  
 Db 594 TTCTGATGATTAATGAAGAACATGATGCGGAAGGAGGCTCTGGGTTAACTGACGCG 653  
 Qy 726 TCATGACGAAAGCGTGGGAGCAACAGATTTAGTACCTTGTATGCTCACGCCCTTAA 785  
 Db 654 TGAAGTGCAGAAAGCGTGGGAGCAACAGATTTAGTACCTTGTATGCTCACGCCCTTAA 713  
 Qy 786 CGATGCTCACTAGTTGTTGGGCTTATTAGGCTTGTGTAAGAACTTAACGCTGAAGTTG 845  
 Db 714 CGATGAATTTAGGCGCTCGGCGAGTTTACTGCTGCTGCGGAGCTTAACGCTTAACAT 773  
 Qy 846 ACCGCTGGGAGTACGCTGCGCAAGTTAACTCAAAAGAAATGACGAGGAGCCGACACA 905  
 Db 774 TCGGCTGGGAGTACGCTGCGCAAGTTAACTCAAAAGAAATGACGAGGAGCCGACACA 833  
 Qy 906 AGCGGTGATTAATGAGTAAATTCGATGCAACGCAAAACCTTACTTACCTTGAAT 965  
 Db 834 AGCGGTGATTAATGAGTAAATTCGATGCAACGCAAAACCTTACTTACCTTGAAT 893  
 Qy 966 GTAGCAATTTTCTAGAGA-----TAGATTAAGTCTTGGGAAAGCTTAACAGTGTGCTG 1020  
 Db 894 CCGGATCGGGAATACGAGAGATGATCTTCTGAGTTCCGCTGATGCGAGACGAGTCTG 953  
 Qy 1021 CATGCTGTGCTGATGCTGCTGCTGAGATTTGGTGTGTAAGTCCGCAACGAGCGCAAC 1080  
 Db 954 CATGCTGTGCTGATGCTGCTGCTGAGATTTGGTGTGTAAGTCCGCAACGAGCGCAAC 1013  
 Qy 1081 CTGTGATTAATGATGATCATCTTGTGTTGGCACTTAAATGAGACTGCGGTGACAAACG 1140  
 Db 1014 CTGCGCTTATGTTGCGAGCATTCAGTTGGGCACTTAAGGAGACTGCGGTGATTAAGCGG 1073



RESULT 61

US-11-055-637-74  
 ; Sequence 74, Application US/11055637  
 ; Publication No. US20050260619A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BROUSSEAU, Roland  
 ; APPLICANT: DUBOIS, Jason  
 ; APPLICANT: EDGE, Tom  
 ; APPLICANT: MASSON, Luc  
 ; APPLICANT: TREVORS, Jack T.  
 ; TITLE OF INVENTION: A DNA MICROARRAY FOR FINGERPRINTING AND  
 ; FILE OF INVENTION: CHARACTERIZATION OF MICROORGANISMS IN MICROBIAL COMMUNITIES  
 ; FILE REFERENCE: 2139-33US  
 ; CURRENT APPLICATION NUMBER: US/11/055,637  
 ; CURRENT FILING DATE: 2005-02-11  
 ; PRIOR APPLICATION NUMBER: US 60/543,288  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 74  
 ; LENGTH: 1535  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Probe for DNA array  
 US-11-055-637-74

Query Match 56.8%; Score 827; DB 17; Length 1535;  
 Best Local Similarity 75.1%; Pred. No. 3.5e-248;  
 Matches 1104; Conservative 0; Mismatches 350; Indels 17; Gaps 5;

4 GAACGCTGGCGGCAAGCTTATACATGCAAGTGAACGGGAGAGCGAGATGCTGATCTG 63  
 24 GAACGCTGGCGGCAAGCTTATACATGCAAGTGAACGGGAGATGCTGATCTG 83  
 64 GTGGCAGTGGCGGCAAGCTTATACATGCAAGTGAACGGGAGATGCTGATCTG 122  
 84 GAAGTTAGCGGCGGCAAGCTTATACATGCAAGTGAACGGGAGATGCTGATCTG 143  
 123 GCATCGAAGATGCTTATATACCGCATTA-----TACTCTAAGAGAGAAAGCGAGG 173  
 144 TTCGGGAAACCGGAGCTTATATACCGGATTACTTCTTCTCTCGAGAAAGGTTGAAAG 203  
 174 ATCGAAGACCTTGGCTTTTGAAGCGCGGATGCTGATGCTGATGCTGATGCTG 223  
 204 ACGGCTTGGCTTCTCACTTACAGATGGCGCGGCGCATTAAGTGTGAGGTAC 263  
 234 GGCCTACCAAGCGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 293  
 264 GGCCTACCAAGCGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 323  
 294 AGACACGGCCAGACTTCTACGGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGG 353  
 324 AGACACGGCCAGACTTCTACGGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGG 383  
 354 CCGATTCACGATGCTGCTGAGTGAAGAGGCTTCCGGTTGTAAGTCTTTGAGTC 413  
 384 TCTGACGAGCAACGCGCGTGAAGTGAAGGTTTCCGATGTAAGTCTTTGATCA 443  
 414 GAGAGAAAGGTTACGCTTAATTAATGATGCTGATGCTGATGCTGATGCTGATG 473  
 444 GAGAGAAAGGTTACGCTTAATTAATGATGCTGATGCTGATGCTGATGCTGATG 503  
 474 GGCCTACCAAGCGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 533  
 504 GGCCTACCAAGCGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 563  
 534 TGGCGCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 593  
 564 TGGCGCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 623  
 594 TGGGATTTGGCTTGAAGTCAAGGCTAGAGTGGCAGAGGAGGAGGAGGAGGAGG 653

624 GTGGAAGGTCATTGGAAGCTGGAGAGCTTGAAGTGAAGAGAGAGGAGGATTCACAGT 683  
 654 GTAGCAGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 713  
 684 GTAGCAGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 743  
 714 TAACTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773  
 744 TGTAACTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803  
 774 CCAGCCCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831  
 804 CCAGCCCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 863  
 832 AACGCTGAAGTGAACGCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 891  
 864 AACGCTGAAGTGAACGCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 923  
 892 CGGGGACCGGACCAAGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 951  
 924 CGGGGACCGGACCAAGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 983  
 952 CCTACCTTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008  
 984 CAGGCTTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1043  
 1009 ACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068  
 1044 TGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1103  
 1069 ACAGCGCAACCTTGTCTATTAATGATGATGATGATGATGATGATGATGATGATGAT 1128  
 1104 ACAGCGCAACCTTGTCTATTAATGATGATGATGATGATGATGATGATGATGATGAT 1163  
 1129 GGTGCAACCGGAGAGGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1188  
 1164 GGTGCAACCGGAGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1223  
 1189 CTTCACGCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1248  
 1224 CTACACGCTGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1283  
 1249 TGAGAAAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308  
 1284 CATATAACCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1343  
 1309 GCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1368  
 1344 GCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1403  
 1369 CCGTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426  
 1404 CCGTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1463  
 1427 GCTTGCACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
 1464 GCCGCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1494

RESULT 62

US-09-941-095-160  
 ; Sequence 160, Application US/09941095  
 ; Publication No. US20060068383A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BROW, MARY ANN D.  
 ; APPLICANT: LYANCHIEV, VICTOR I.  
 ; APPLICANT: OLIVE, DAVID M.  
 ; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
 ; PATHOGENS  
 ; NUMBER OF SEQUENCES: 165  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MEDLEN & CARROLL

STREET: 220 MONTGOMERY STREET, SUITE 2200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/941,095  
FILING DATE: 28-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/941,193  
FILING DATE: 28-Aug-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, PETER G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: FONS-01756  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 160:  
US-09-941-095-160

Query Match 56.7%; Score 825.6; DB 8; Length 1555;  
Beet Local Similarity 75.8%; Pred. No. 9,7e-248;  
Matches 1116; Conservative 0; Mismatches 339; Indels 17; Gaps 7;

QY 3 TGAAGCGTGGCGGATGCTTTACATGCAAGTCCGACGCGACAG-GATGCTTGATC 61  
DB 31 TGAAGCGTGGCGGATGCTTTACATGCAAGTCCGACGCGACAGCGAAGACCTTGCTTC 90  
QY 62 T-GGTGGCGAGTGGCGGACGGGTGAGTAATGATCGG-AACTATCCAGAGAGGGGGT 119  
DB 91 TCTATGTTAGCGGCGAGCGGTGAGTAAACAGTGAATACCTTAATGAAGTGGGAT 150  
QY 120 AACGCATGAAAGATGTGTCTAATACCGCAT------TACTCTAAGGAGAAAGCAG 171  
DB 151 AACTTCGGGAAACCGAGCTAATACCGGATATATTTGAAACCGCATGCTTCAAAAGTGA 210  
QY 172 GGAATGAAAGACCTTGCGCTTTTGAGAGCGGACGATGTCTGAATTAAGTGTGGGGA 231  
DB 211 AAGAGGCTTGTCTGCTAATATAGATGATCCGCTGCAATTAAGTGTGAAGTA 270  
QY 232 AAGGCTTACCAAGGCGAGATGATGATGCTGAGAGAGCAGACCACTGAGAC 291  
DB 271 ACGGCTTACCAAGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 330  
QY 292 TGAAGCAGCGGCGAGATGCTTAAGGAGGAGCAGATGGGGAATTTTGAACAATGGGCGCA 351  
DB 331 TGAAGCAGCGGCGAGATGCTTAAGGAGGAGCAGATGGGGAATTTTGAACAATGGGCGCA 390  
QY 352 AGCTGATCCAGCAATGCGCGGTGAGTGAAGAGGCTTTGGGTTGTAAAGTCTTTGAG 411  
DB 391 AGCTGATCCAGCAATGCGCGGTGAGTGAAGAGGCTTTGGGTTGTAAAGTCTTTGAG 450  
QY 412 TCGAGAGAAAGGTTACGTTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 471  
DB 451 TAGGAGAAACATATGTGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510  
QY 472 CCGGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531  
DB 511 ACGGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570

QY 532 ACTGGCGCTAAAGGCTGCGAGCGGCTTTGTAAGTCAATGATGTAATCCCGGCTTAA 591  
DB 571 ATTGGCTTAAGGCGGCTGAGCGGCTTTTAAAGTCAATGATGTAATCCCGGCTTAA 630  
QY 592 CCGGAGATGCTGTTTGAACAATGAGTGTGCGAGAGAGGAGTGAATTCAT 651  
DB 631 CCGGAGATGCTGTTTGAACAATGAGTGTGCGAGAGAGGAGTGAATTCAT 690  
QY 652 GTGAGCAGTGAATGCTGAGATGAGTGAAGAACATGATGAGGAGGAGGCTCTCG 711  
DB 691 GTGAGCAGTGAATGCTGAGATGAGTGAAGAACATGATGAGGAGGAGGCTCTCG 750  
QY 712 GTTAACTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771  
DB 751 TCTGTAATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 810  
QY 772 GTTCAAGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829  
DB 811 GTTCAAGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870  
QY 830 CTAAAGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 889  
DB 871 CTAAAGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930  
QY 890 GACGGGAGCGGCGACCAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 949  
DB 931 GACGGGAGCGGCGACCAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 990  
QY 950 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1006  
DB 991 TACCTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050  
QY 1007 TAAACAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066  
DB 1051 AGTGAAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1110  
QY 1067 CAAGAGCGCAACCTTGTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1126  
DB 1111 CAAGAGCGCAACCTTGTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1170  
QY 1127 CCGGTGACAAACCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186  
DB 1171 CCGGTGACAAACCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1230  
QY 1187 GCTTCAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246  
DB 1231 GCTTCAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290  
QY 1247 TCTCAGAAAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306  
DB 1291 TCTCAGAAAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350  
QY 1307 TCGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366  
DB 1351 TCGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1410  
QY 1367 GCGGCTCAGCAGTGAAGTGGTTTCAACGAGAGCAGTGTGTAAC-CTAAGAGGG 1425  
DB 1411 GCGGCTCAGCAGTGAAGTGGTTTCAACGAGAGCAGTGTGTAAC-CTAAGAGGG 1470  
QY 1426 GCGTTCAGCAGTGAAGTGAATTCATGCTGGGGT 1457  
DB 1471 AGCGGTGAGAGTGGAGCAATGATGGGGT 1502

RESULT 63  
US-11-198-746-160  
; Sequence 160, Application US/11198746  
; Publication No. US2006035256A1  
; GENERAL INFORMATION:  
; APPLICANT: BROW, MARY ANN D.  
; APPLICANT: LYAMICHEV, VICTOR I.

APPLICANT: OLIVE, DAVID M.  
TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
TITLE OF INVENTION: PATHOGENS  
NUMBER OF SEQUENCES: 165  
CORRESPONDENCE ADDRESS:  
ADDRESSER: MEDLEN & CARROLL  
STREET: 220 MONTGOMERY STREET, SUITE 2200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/11/198,746  
FILING DATE: 05-Aug-2005  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/09/941,193  
FILING DATE: 28-Aug-2001  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, PETER G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: FORS-01756  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-11-198-746-160

Query Match 56.7%; Score 825.6; DB 18; Length 1555;  
Best local Similarity 75.8%; Pred. No. 9,7e-248;

Matches 1116; Conservative 0; Mismatches 339; Indels 17; Gaps 7;

3 TGAAGCTGCGCGGCGCTTACATGCAAGTCGAGCGACGACGCG-ATGCTTCATC 61  
1 TGAAGCTGCGCGGCGCTTACATGCAAGTCGAGCGACGACGAGGAGCTTCTTC 90  
62 T-GGTGCGAGTGGCGGCGGAGTGAATGATCGG-AACTATCCAGAGAGGGGGT 119  
91 TCTGATGTTAGCGCGCGAGCGGTGATGATCAAGTGAATACCTATGATGATGATGAT 150  
120 AACGATCGAAGATGCTTAATACCGCAT------TACTCTAAGAGAGAAAGCAG 171  
151 AACTTGGGGAACCGGAGCTAATATACGGAATATTTTGAACCGCATGCTTCAAAAGTGA 210  
172 GGAATCGAAGACCTTGCGCTTTTGGAGCGCGCGATGCTGATAGCTAGTGTGGGGTGA 231  
211 AAGACGCTCTTGTCTACTTATGATGATGATCGCGCTGCTTATGATGATGATGATGAT 270  
232 AAGGCTTACCAAGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 291  
271 ACGGCTTACCAAGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 330  
292 TGAAGACGCGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351  
331 TGAAGACGCGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 390  
352 AGGCTTACCAAGCGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 411  
391 AGGCTTACCAAGCGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450  
412 TCAGAGAAAGAGGTTACGTTAATATCGTGAATGATGATGATGATGATGATGATGATGAT 471

451 TAGGAGAAACATATGTGTAGTACTGACATCTTGAAGGATCTTAATCAGAAAGCC 510  
472 CCGGCTTACATGATGCGCAGACCGCGGTAATACGATGATGATGATGATGATGATGATGAT 531  
511 ACGGCTTACATGATGCGCAGACCGCGGTAATACGATGATGATGATGATGATGATGATGAT 570  
532 ACTGGGCGTAAAGGATGCGGCGGCTTTGTAATGATGATGATGATGATGATGATGATGAT 591  
571 ATTGGGCGTAAAGGATGCGGCGGCTTTTGAATGATGATGATGATGATGATGATGATGAT 630  
592 CCGGGAATGCGGCTTGAATCTAACGAGCTAGTGTGCGAGGAGGATGATGATGATGATGAT 651  
631 CCGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690  
652 GTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711  
691 GTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 750  
712 GTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771  
751 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 810  
772 GTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829  
811 GTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870  
830 GTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 889  
871 GTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930  
890 GATGGGGAACCGCGCAGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 949  
931 GATGGGGAACCGCGCAGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 990  
950 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1006  
991 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050  
1007 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066  
1051 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1110  
1067 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126  
1111 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1170  
1127 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186  
1171 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1230  
1187 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246  
1231 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290  
1247 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306  
1291 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350  
1307 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366  
1351 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1410  
1367 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1425  
1411 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470  
1426 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
1471 TACCTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1502

RESULT 64



Qy 1426 CGCTTGCCAGCGTGAGATTGACTGAGGCTG 1457  
Db 1471 AGCCGTGCAAGGTGGGACAAATGATTGGGGTG 1502

RESULT 65  
US-11-198-657-160  
Sequence 160, Application US/11198657  
Publication No. US20060040299A1

GENERAL INFORMATION:

APPLICANT: BROW, MARY ANN D.  
APPLICANT: LYAMICHEV, VICTOR I.  
APPLICANT: OLIVE, DAVID M.  
TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
TITLE OF INVENTION: PATHOGENS  
NUMBER OF SEQUENCES: 165  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MEDLEN & CARROLL  
STREET: 220 MONTGOMERY STREET, SUITE 2200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 05-Aug-2005  
APPLICATION NUMBER: US/11/198,657  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/941,193  
FILING DATE: 28-Aug-2001  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, PETER G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: FORS-01756  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)

US-11-198-657-160

Query Match 56.7%; Score 825.6; DB 18; Length 1555;  
Best Local Similarity 75.8%; Pred. No. 9.7e-248;  
Matches 1116; Conservative 0; Mismatches 339; Indels 17; Gaps 7;

Qy 3 TGAACGTGCGGCGCATGCTTTACATACATGCAAGTCGAACGCGACGACG-GATGCTTGATC 61  
Db 31 TGAACGTGCGGCGCGCTGCTTAATACATGCAAGTCGAGCGAAGCGAAGAGGTTGCTTC 90  
Qy 62 T-GGTGGCGAGTGGCGGACGCGGTGAGTATGATCGG-ANCGTATCCAGAAAGGGGGGT 119  
Db 91 TGTGATGTGAGCGGCGGACGCGGTGAGTATACCGTGTGATACCTATTAAGACTGGGAT 150  
Qy 120 AAGCATCGAAGATGTGCTAATACCGCAT------TACTCTAAGAGGAAGCAGG 171  
Db 151 AACTTCGGGAAACCGGAGCTAATACCGGATTAATTTTGAACCGCATGTTCAAAAGTGA 210  
Qy 172 GGATCGAAGACCTTGGGCTTTTGGAGCGGCGGATGCTGATTAGCTAGTGTGGGCTA 231  
Db 211 AAGACGCTTCTGCTGCTAATTAAGATGATCCGCGCTGATTAAGTAAAGTGA 270  
Qy 232 AAGGCTTACCAAGCGGACGATCACTAGTGTGCTGAGAGACGACGACCACTGGGAC 291

Db 271 AAGGCTTACCAAGCGGACGATCACTAGTGTGCTGAGAGGGGTGATCGGCCACTGGAAC 330  
Qy 292 TGAACACGCGCCGAGCTTCTTAACGAGGACGAGCTGGGGAATTTTGGACAAATGGGCGCA 351  
Db 331 TGAACACGCTTCAGACTCTTACCGGAGGACGAGTAAATCTTCCGCAATGGGCGCA 390  
Qy 352 AGCGTATCCGACATGCGCGGTAGTGAAGAAAGCCCTCGGGTGTAAAGCTCTTTGAG 411  
Db 391 AGCGTATCCGACATGCGCGGTAGTGAAGAAAGCCCTCGGGTGTAAAGCTCTTTGAG 450  
Qy 412 TCGAAGAAAGAAAGTTACGCTAAATATCGTACTCATGACGCTATCGACAGAAAGCA 471  
Db 451 TAGGGAAGAAATATGTAAGTAACTGTGACATCTTGAACGCTAATCAGAAAGCC 510  
Qy 472 CCGGCTAATACGTGCGACGACCGCGGTATATGATGAGGTGCAAGCCTTAATGGAATT 531  
Db 511 ACGGCTAATACGTGCGACGACCGCGGTATATGATGAGGTGCAAGCCTTAATGGAATT 570  
Qy 532 ACTGGCGCTAAAGGTGGCGGCGGCTTTGTAAGTCAATGTGAAATCCCGGCTTAA 591  
Db 571 ATTGGCGTAAAGCGCGGTGAGCGGCTTTTAAAGTCAATGTGAAATCCCGGCTTAA 630  
Qy 592 CCGTGAAGGTGCTTTGAACTACAAAGGCTAAGTGTGACAGAGGAGGTGGAATTCAT 651  
Db 631 CCGTGAAGGTGCTTTGAACTACAAAGGCTAAGTGTGACAGAGGAGGTGGAATTCAT 690  
Qy 652 GTTACGAGTGAATGCTGATGATATGGAAGAACTGATGCGCAAGGAGGCTCTCGG 711  
Db 691 GTTACGAGTGAATGCTGATGATATGGAAGAACTGATGCGCAAGGAGGCTCTCTCGG 750  
Qy 712 GTTACGAGTGAATGCTGATGATATGGAAGAACTGATGCGCAAGGAGGCTCTCTCGG 771  
Db 751 TGTGTAAGTGAATGCTGATGATATGGAAGAACTGATGCGCAAGGAGGCTCTCTCGG 810  
Qy 772 GTTACGAGTGAATGCTGATGATATGGAAGAACTGATGCGCAAGGAGGCTCTCTCGG 829  
Db 811 GTTACGAGTGAATGCTGATGATATGGAAGAACTGATGCGCAAGGAGGCTCTCTCGG 870  
Qy 830 GTTACGAGTGAATGCTGATGATATGGAAGAACTGATGCGCAAGGAGGCTCTCTCGG 889  
Db 871 GTTACGAGTGAATGCTGATGATATGGAAGAACTGATGCGCAAGGAGGCTCTCTCGG 930  
Qy 890 GACGCGGACCGCGCAAGGCGGTGATTAATGATGATTAATGATGCAACGCAAAACCT 949  
Db 931 GACGCGGACCGCGCAAGGCGGTGATTAATGATGATTAATGATGCAACGCAAAACCT 990  
Qy 950 TACCTACCTGACATGATGCAATTTTCTAGAGATGA---TTAGTGTCTGGGAAAGC 1006  
Db 991 TACCTACCTGACATGATGCAATTTTCTAGAGATGA---TTAGTGTCTGGGAAAGC 1050  
Qy 1007 TAAACAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1066  
Db 1051 AGTGAAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1110  
Qy 1067 CAACGAGCGCAACCTTGTGATTAATGATGATTAATGATGATTAATGATGATGACTG 1126  
Db 1111 CAACGAGCGCAACCTTGTGATTAATGATGATTAATGATGATTAATGATGATGACTG 1170  
Qy 1127 CCGGTGACAAACCGGAGAAAGGTGGGATGAGTCAAGTCTCAATGCGCTTATGGGTAG 1186  
Db 1171 CCGGTGACAAACCGGAGAAAGGTGGGATGAGTCAAGTCTCAATGCGCTTATGGGTAG 1230  
Qy 1187 GGGTTCACAGTAAATGCAATGCGCGGTACAGAGGTTTGCACACCGCGAGGGGAGCTTA 1246  
Db 1231 GGGTTCACAGTAAATGCAATGCGCGGTACAGAGGTTTGCACACCGCGAGGGGAGCTTA 1290  
Qy 1247 TCTCAGAAAGCGCGTGTGATGCTCGATGCGAAGTCTCAATGCGCTTATGGGTAG 1306  
Db 1291 TCTCAGAAAGCGCGTGTGATGCTCGATGCGAAGTCTCAATGCGCTTATGGGTAG 1350  
Qy 1307 TCGCTAGTATCGCGGATGAGATGCGCGGTAAATCGTTCCCGGGTCTTGTACACACC 1366  
Db 1351 TCGCTAGTATCGCGGATGAGATGCGCGGTAAATCGTTCCCGGGTCTTGTACACACC 1410



Oy 1367 GCCGCTCACACCATGGAGTGGGTTTCACACAGACAGGTAGTCTTAACC-GTAAGAGAGG 1425  
Db 1411 GCCCCTTCACACCAACAGAGTTTGTAAACACCCGAAACCCGCTGAGTAACCTTTTAGAGACT 1470  
Oy 1426 CGCTTGCCACGGTGAATTTCATGACTGGGGTG 1457  
Db 1471 AGCCGTCGAAGTGGGACAAATGATTGGGGTG 1502

## RESULT 66

US-11-084-508-3  
; Sequence 3, Application US/11084508  
; Publication No. US20050260737A1  
; GENERAL INFORMATION:  
; APPLICANT: Rahman, Raja Noor Zailha Abd.  
; APPLICANT: Salleh, Abu Bakar  
; APPLICANT: Basri, Mahran  
; APPLICANT: Hun, Chin John  
; TITLE OF INVENTION: Novel Lipase Gene from *Bacillus sphaericus* 205y  
; FILE REFERENCE: KAN-101  
; CURRENT APPLICATION NUMBER: US/11/084,508  
; PRIOR FILING DATE: 2005-03-18  
; PRIOR APPLICATION NUMBER: MY 20040958  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 1459  
; TYPE: DNA  
; ORGANISM: *Bacillus sphaericus*  
US-11-084-508-3

Query Match 56.5%; Score 823.4; DB 17; Length 1459;  
Beet Local Similarity 76.4%; Pred. No. 4,7e-247;  
Matches 1093; Conservative 0; Mismatches 321; Indels 17; Gaps 6;

Oy 4 GAACGCTGGCGGCATGCTTTACATGCAAGTCGAACG-GCAGCAACGATGCTTCATC- 61  
Db 29 GAACCTGCGCGCGTCCCTTAATACATGCAAGTCGAACGAAAGAGCTTCCT 88  
Oy 62 TGGTGGCAGTGGCGGAGCGGTGAATATGATCCGG-AACTATCCAGAAAGAGGGGTA 120  
Db 89 TTGACGTTAAGCGCGGAGCGGTGAATACATGCGGCAACTTAATAGTTGGGATA 148  
Oy 121 ACGCATCAAGAAAGTGTCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAA 180  
Db 149 ACTCGGGAAACCGGGGCTAATACGAATATCTATTTCATTGGAATTACTGAA 208  
Oy 181 GAC-----CTTGGCGTTTGGAGCGGCGGATGTCTGATTACTGTTGGTGGGTA 231  
Db 209 AGACGCTCTCGGCTGTCTATMAAGTGGCCGCGGCATTAAGTGTAGGTA 268  
Oy 232 AAGGCTTACAAAGGCGAGCATAGTGTCTGAGAGAGCAACAGCACTGGAGC 291  
Db 269 ACGGCTTACAAAGGCGAGCATAGTGTCTGAGAGGATCGGCGCACTGGGAC 328  
Oy 292 TGAAGACAGCGCCAGACTTCTTACGAGGAGCAGCAGTGGGAAATTTTGAACAATGGCGCA 351  
Db 329 TGAGACAGCGCCAGACTTCTTACGAGGAGCAGCAGTGGGAAATTTTCCCAATGGCGCA 388  
Oy 352 AGCCTGATCCAGCAATGCCGCTGAGTGAAGAAAGCCTTCGGGTTGTAAGCTCTTTGAG 411  
Db 389 AGCCTGATGAGCAACGCGCGTGAAGTGAAGAAAGGTTTTCGATGTAAGCTCTGTTGT 448  
Oy 412 TCGAGAGAAAGGTTACGTTAATATGTAATCTGACTCATGACGGTATTCGACAGAAAGACA 471  
Db 449 AAGGAAAGAAAGTACAGTATGTAATCTGCTTACTTGAACGTTACTTTATTAAGAAACC 508  
Oy 472 CCGGCTTAACTAGTCCAGCAGCAGCGCGTGAATATGTAAGGTTGCAAGCGTTAATCGAAAT 531  
Db 509 ACGGCTTAACTAGTCCAGCAGCAGCGCGGTTAATATGTAAGGTTGCAAGCGTTTGTCCGAAAT 568

Oy 532 ACTGAGCGTAAAGGGGTGGCGAGCGGCTTTTGAAGTCAGATGTGAATATCCCGGGCTTAA 591  
Db 569 ATTGGCGTAAAGGGCGCGCGGCTCTTTAAGTCTGATGTGAATATCCCGGGCTTAA 628  
Oy 592 CTTGGGAATTTGCGTTTGAATCTAAGAGTGTGGCAGAGGAGTGAATTTCCAT 651  
Db 629 CCGTGAAGGTCATTTGGAATCTGGGGGACTTGAAGTGAAGAAAGTGAATTTCCAA 688  
Oy 652 GTTAGCAGTGAATTTGGGTGAAGTATGGAAGAAATCATGATGGGAAAGGCGGCTCTGG 711  
Db 689 GTTAGCGGTGAATTTGGGTGAAGTATGGAAGAAATCATGATGGGAAAGGCGGCTCTGG 748  
Oy 712 GTTAACACTGACGCTCATGACGAAAGCGTGGGAGCAACGAGATTAGATCCCTGTA 771  
Db 749 TCTGTAATGACGCTGAGCGCGCGGAAACGCTGGGAGCAACGAGATTAGATCCCTGTA 808  
Oy 772 GTTCAGCGCTTAAACGATGTCAATGATGTTT-GGGCTTATTAAGCTTGTGAACGAG 829  
Db 809 GTTCAGCGCTTAAACGATGTCAATGATGTTTGGGGGTTTCCCGCTTATGCTGTCAG 868  
Oy 830 CTAACGCGTGAATTTGACCGCTTGGGAGTACCGTCCCAAGTTAAATCAAGAAAT 889  
Db 869 CTAACGCAATTAAGCACTCCGCTGGGAGTACCGTCCCAAGTGAATCTTAAGAAAT 928  
Oy 890 GACGGGAGCCCGCAACGCGGTGATTTATGATTAATTCATGATGCAACGCAAAACCT 949  
Db 929 GACGGGAGCCCGCAACGCGGTGATTTATGATTAATTCATGATGCAACGCAAAACCT 988  
Oy 950 TACCTACCTTGAAT--GTAGCAATTTTCTAAGATGATTAATGCTTCGGGAAACG 1006  
Db 989 TACCAAGCTTGAATCCCGTTGACCACTGAATGATTAATGCTTCCTTCGGGAAACG 1048  
Oy 1007 TAAACAGTGTCTCAATGCTGTCTGACGCTCGTGTGAGATGTTGGGTTAGTCCG 1066  
Db 1049 GGTGACAGTGTCTCAATGCTGTCTGACGCTCGTGTGAGATGTTGGGTTAGTCCG 1108  
Oy 1067 CAACGAGCGCAACCTTGTCAATTAATGCTCAATTTGTTGGGCACTTAATGAGACTG 1126  
Db 1109 CAACGAGCGCAACCTTGTCAATTAATGCTCAATTTGTTGGGCACTTAATGAGACTG 1168  
Oy 1127 CCGGTGAACAAACCGGAGAGTGGGATGACGTCAGTCTCAATGCGCTTATGGGTAG 1186  
Db 1169 CCGGTGAACAAACCGGAGAGTGGGATGACGTCAGTCTCAATGCGCTTATGAGACTG 1228  
Oy 1187 GCTTCAACGTAATTAATTAATGCGCGCTGACAGAGGTTGGCAACCGGAGGAGGACTAA 1246  
Db 1229 GCTTCAACGTAATTAATTAATGCGCGCTGACAGAGGTTGGCAACCGGAGGAGGACTAA 1288  
Oy 1247 TCTCAGAAAGCGGCTGTAATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1306  
Db 1289 TCGATTAAGTGTCTCAATGCTGATTTGAGGCTGCACTCGGCTTATGAGAAACCGGAA 1348  
Oy 1307 TCGCTAATTAATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1366  
Db 1349 TCGCTAATTAATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1408  
Oy 1367 GCCCGTCAACCAATGGAAGTGGTTTCAACGAAAGAGTGAATCCG 1417  
Db 1409 GCCCGTCAACCAATGGAAGTGGTTTCAACCGAAAGTGGTGGGTAACCTT 1459

## RESULT 67

US-11-045-004-1  
; Sequence 1, Application US/11045004  
; Publication No. US20060079901A1  
; GENERAL INFORMATION:  
; APPLICANT: BUCHRISER, CARMEN  
; APPLICANT: FRANGUL, LIONEL  
; APPLICANT: COUVE, ELISABETH  
; APPLICANT: RUSNIOK, CHRISTOPHE  
; APPLICANT: FSIHI, HAIDA  
; APPLICANT: DEHOUC, PIERRE  
; APPLICANT: DUSSURGET, OLIVIER



```

APPLICANT: CHETOUANI, PARID
APPLICANT: NEDJARI, HAFED
APPLICANT: GLASER, PHILIPPE
APPLICANT: KUNST, FRANCK
APPLICANT: COSSART, PASCALE
APPLICANT: DANIELS, JUSTIN
APPLICANT: GOEBEL, WERNER
APPLICANT: KREFT, JURGEN
APPLICANT: KUHN, MICHAEL
APPLICANT: NG, EVA
APPLICANT: VAZQUEZ-BOLAND, ANTONIO
APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
APPLICANT: GARRIDO-GARCIA, PATRICIA
APPLICANT: TIERREZ-MARTINEZ, ALBERTO
APPLICANT: AMEND, ALEXANDRA
APPLICANT: CHAKRABORTY, TRINDAD
APPLICANT: DOMANN, EUGEN
APPLICANT: HAIN, THORSTEN
APPLICANT: BERCHE, PATRICK
APPLICANT: CHARBIT, ALAIN
APPLICANT: DURANT, LIONEL
APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
APPLICANT: BAQUERO, FERNANDO
APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
APPLICANT: GOMEZ-LOPEZ, NURIA
APPLICANT: MADUENIO, ENCARNIA
APPLICANT: PABLOS, BETRIZ DE
APPLICANT: MEHLAND, JURGEN
APPLICANT: KARST, UWE
APPLICANT: ENTJAN, KARL-DIETER
APPLICANT: HAUF, JORG
APPLICANT: ROSE, MATTHIAS
APPLICANT: VOSS, HAMUT
APPLICANT: TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
APPLICANT: FILE REFERENCE: 05394.0018-02
APPLICANT: CURRENT APPLICATION NUMBER: US/11/045,004
APPLICANT: PRIOR FILING DATE: 2005-01-28
APPLICANT: PRIOR APPLICATION NUMBER: 10/637,657
APPLICANT: PRIOR FILING DATE: 2003-08-11
APPLICANT: PRIOR APPLICATION NUMBER: 10/257,023
APPLICANT: PRIOR FILING DATE: 2002-10-08
APPLICANT: PRIOR APPLICATION NUMBER: PCT/FR01/01118
APPLICANT: PRIOR FILING DATE: 2001-04-11
APPLICANT: PRIOR APPLICATION NUMBER: FR 00/04,629
APPLICANT: PRIOR FILING DATE: 2000-04-11
APPLICANT: NUMBER OF SEQ ID NOS: 2854
APPLICANT: SOFTWARE: PatentIn version 3.3
APPLICANT: SEQ ID NO 1
APPLICANT: LENGTH: 2944528
APPLICANT: TYPE: DNA
APPLICANT: ORGANISM: Listeria monocytogenes
US-11-045-004-1

Query Match 56.5%; Score 823.2; DB 18; Length 2944528;
Best Local Similarity 76.2%; Pred. No. 3.5e-246;
Matches 1120; Conservative 0; Mismatches 333; Indels 17; Gaps 8;

```

```

QY 232 AAGGCTACCAAGGCGAGATCAATGTTGGTCTGAGAGAGACAGACCACTACTGGGNC 291
DB 237739 ATGGCTTACCAAGGCGAGATCAATGTTGGTCTGAGAGAGAGACAGACCACTACTGGGNC 237798
QY 292 TGAGACAGGCGCCAGACTCTTACCGGAGGAGCAGCACTGGGGAATTTTGAACAAATGGGCGCA 351
DB 237739 TGAGACAGGCGCCAGACTCTTACCGGAGGAGCAGCACTGGGGAATTTTCCGCAATGAGCGAA 237858
QY 352 AGCTTATCCAGCAATGCCGCGGTGATGTAAGAAAGGCCCTTCGGGTGTGAAGCTCTTTGAG 411
DB 237859 AGCTTATCCAGCAATGCCGCGGTGATGTAAGAAAGGCCCTTCGGGTGTGAAGCTCTTTGAG 237918
QY 412 TCGAAGAAAGGTTAAGGTTAATATCGTACTCATGACCGTATCGACAGAAAGCA 471
DB 237919 TCGAAGAAAGGTTAAGGTTAATATCGTACTCATGACCGTATCGACAGAAAGCA 237978
QY 472 CCGGCTTACTACGTCAGCAGACCGCGGTATATACGTAAGGTCAGACCGTTAATCGGAAT 531
DB 237979 ACGGCTTACTACGTCAGCAGACCGCGGTATATACGTAAGGTCAGACCGTTAATCGGAAT 238038
QY 532 ACTGGGCGTAAAGGTCGCGAGCGGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAA 591
DB 238039 ACTGGGCGTAAAGGTCGCGAGCGGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAA 238098
QY 592 CCGGGAATTCGCTTGAACCTAACAGGCTAGAGTGGCAGAGGAGGTGGAATTCAT 651
DB 238099 CCGGGAATTCGCTTGAACCTAACAGGCTAGAGTGGCAGAGGAGGTGGAATTCAT 238158
QY 652 GTGTAGCAGTGAATATCGTGAAGATATGGAAGAACATCGATGGCGAGAGCGCTCTGG 711
DB 238159 GTGTAGCAGTGAATATCGTGAAGATATGGAAGAACATCGATGGCGAGAGCGCTCTGG 238218
QY 712 GTTAACTAGCAGCTCATACAGAAAGCGTGGGAGCAACAGAGTAAATACCTCTGTA 771
DB 238219 GTTAACTAGCAGCTCATACAGAAAGCGTGGGAGCAACAGAGTAAATACCTCTGTA 238278
QY 772 GTTCAAGCCCTTAAAGATGTCATCTAGTTGT - GGGCGCTTATAGGCTTGTGAAG 829
DB 238279 GTTCAAGCCCTTAAAGATGTCATCTAGTTGT - GGGCGCTTATAGGCTTGTGAAG 238338
QY 830 CTAAAGCGTGAAGTTGACCGCTGCGGAGTACGCTGCAAGATTAATCAAGAGAT 889
DB 238339 CTAAAGCGTGAAGTTGACCGCTGCGGAGTACGCTGCAAGATTAATCAAGAGAT 238398
QY 890 GACGGGAGCCCGCAGCAAGCGGTGATATGAGATTAATTCATGCAAGCGCAAAACCT 949
DB 238399 GACGGGAGCCCGCAGCAAGCGGTGATATGAGATTAATTCATGCAAGCGCAAAACCT 238458
QY 950 TACCTTACCTTGAATGACGCAATTTTCTAAGATAGA - TTAGTGCTTCGGGAAGCGTA 1008
DB 238459 TACCTTACCTTGAATGACGCAATTTTCTAAGATAGA - TTAGTGCTTCGGGAAGCGTA 238518
QY 1009 ACAGAGTGTGATGATGCTGTGCTGAGCTCGTGTGAGATGTTGGTTAATGCCGCA 1068
DB 238519 ACAGAGTGTGATGATGCTGTGCTGAGCTCGTGTGAGATGTTGGTTAATGCCGCA 238578
QY 1069 ACAGAGCAACCTCTTGTCAATTAATTCATCAATTTGTTGGGCACTTAAATGAGACTGCC 1128
DB 238579 ACAGAGCAACCTCTTGTCAATTAATTCATCAATTTGTTGGGCACTTAAATGAGACTGCC 238638
QY 1129 GGTGACAAACCGGAGAGGTTGGGATGACGTCAATCTCTATGCGCTTATAGGTAAGG 1188
DB 238639 GGTG - CAAGCCGAGAGAGGTTGGGATGACGTCAATCAATCAATCAATGCCCCCTTATAGCCTGGG 238697
QY 1189 CTTCAACAGTAATACATAGCGCGGTACAAGAGGTTGCAACCGCGAGAGGAGCTAATC 1248
DB 238698 CTTCAACAGTAATACATAGCGCGGTACAAGAGGTTGCAACCGCGAGAGGAGCTAATC 238757
QY 1249 TCAGAAAGCGGTGATGTCGAGTCCGATCGAGTCTGCAACTCGATCCGTGAAGTCGAAATC 1308
DB 238758 TCAGAAAGCGGTGATGTCGAGTCCGATCGAGTCTGCAACTCGATCCGTGAAGTCGAAATC 238817
QY 1309 GCTAGTAATCGCGAGATCAAGATGTCGGCGGTGAATAGCTTCCGGGCTTGTACACACCGC 1368

```

DB 238818 GCTAGTAATCTGTGATCCAGATGCCAGGTGTAATACCTTCCCGGCTTTGTACACACCCG 238877  
QY 1369 CCGTCACACCATGGAGTGGGTTTTCACCAAGAGAGTACTTAACC-GTAAGAGGGCGC 1427  
DB 238978 CCGTCACACCAAGAGTGTGTAACACCCGAAGTGGGTAACTTTATGAGAGCCAG 238937  
QY 1428 CTGGCCACGGTGAATCTATGACTGGGTG 1457  
DB 238938 CCGCCGAAGTGGGACAGATTAATGGGGTG 238967

## RESULT 68

US-11-045-004-1/c  
; Sequence 1, Application US/11045004  
; Publication No. US20060078901A1  
; GENERAL INFORMATION:  
; APPLICANT: BUCHRIESEN, CARMEN  
; APPLICANT: FRANGEUL, LIONEL  
; APPLICANT: COUVE, ELISABETH  
; APPLICANT: RUSNIOK, CHRISTOPHE  
; APPLICANT: FSIHI, HAFLDA  
; APPLICANT: DEHOUX, PIERRE  
; APPLICANT: DUSSENET, OLIVIER  
; APPLICANT: CHETOUNI, FARID  
; APPLICANT: NEDJARI, HAFED  
; APPLICANT: GLASER, PHILIPPE  
; APPLICANT: KUNST, FRANCK  
; APPLICANT: COSSART, PASCALE  
; APPLICANT: DANIELS, JUSTIN  
; APPLICANT: GOEBEL, WERNER  
; APPLICANT: KREFT, JURGEN  
; APPLICANT: KUHN, MICHAEL  
; APPLICANT: NG, EVA  
; APPLICANT: VAZQUEZ-BOLAND, ANTONIO  
; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO  
; APPLICANT: GARRIDO-GARCIA, PATRICIA  
; APPLICANT: TIERREZ-MARTINEZ, ALBERTO  
; APPLICANT: AMEND, ALEXANDRA  
; APPLICANT: CHARRABORTY, TRINAD  
; APPLICANT: DOMANN, EUGEN  
; APPLICANT: HAIN, THORSTEN  
; APPLICANT: BERCHE, PATRICK  
; APPLICANT: CHARBIT, ALAIN  
; APPLICANT: DURANT, LIONEL  
; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO  
; APPLICANT: BAQUERO, FERNANDO  
; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO  
; APPLICANT: GOMEZ-LOPEZ, NURIA  
; APPLICANT: MADUENIO, ENCARN  
; APPLICANT: PABLOS, BETRIZ DE  
; APPLICANT: WEHLAND, JURGEN  
; APPLICANT: KARST, UWE  
; APPLICANT: ENTIAN, KARL-DIETER  
; APPLICANT: HAUF, JORG  
; APPLICANT: ROSE, MATTHIAS  
; APPLICANT: VOSSE, HAMUT  
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME. POLYPEPTIDES AND USES  
; FILE REFERENCE: 05394.0018-02  
; CURRENT APPLICATION NUMBER: US/11/045,004  
; CURRENT FILING DATE: 2005-01-28  
; PRIOR APPLICATION NUMBER: 10/637,657  
; PRIOR FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: 10/257,023  
; PRIOR FILING DATE: 2002-10-08  
; PRIOR APPLICATION NUMBER: PCT/FR01/01118  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: FR 00/04,629  
; PRIOR FILING DATE: 2000-04-11  
; NUMBER OF SEQ ID NOS: 2854  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1  
; LENGTH: 2944528

TYPE: DNA  
; ORGANISM: Listeria monocytogenes  
US-11-045-004-1

Query Match 56.5%; Score 823.2; DB 18; Length 2944528;  
Best Local Similarity 76.2%; Pred. No. 3.5e-246;  
Matches 1120; Conservative 0; Mismatches 333; Indels 17; Gaps 8;

QY 4 GAACGCTGGCGGCATGCTTTACATGCAATCCGAACGCGACGACCGATGCTTGATCT- 62  
DB 1747787 GAACGCTGGCGGCATGCTTTACATGCAATCCGAACGCGACGACCGATGCTTGATCT- 1747728  
QY 63 -GGTGGCAGTGGCGGAGCGGAGTGAATGATCGG- AACGTATCCAGAAAGAGGGGCTA 120  
DB 1747727 CCAAGTTATGTGGCGGAGCGGAGTGAATGATCGG- AACGTATCCAGAAAGAGGGGCTA 1747668  
QY 121 ACGCATCGAAGATGTGCTAATACCGAT- ATACTTAAGAGAGAAACAGGGAGTCGA 179  
DB 1747667 ACTCCGGGAAACCGGGGCTAATACCGAATGATGATGTGGCGCATGCGCTTTTGAA 1747608  
QY 180 AGAC-----CTTGGCCTTTTGAAGCGGCCGATGTCTGATTAGCTAGTTGGGGGTA 231  
DB 1747607 AGATGGTTTCGGCTATGCTTACAGATGGGCCCGCGGTGCTATGCTAGTTGGGGTA 1747548  
QY 232 AAGGCTTACCAAGGCGAGATCAGATGTTGGTCTGAGAGAGCAACACCACTGGGAC 291  
DB 1747547 ATGGCTTACCAAGGCGAGATCAGATGTTGGTCTGAGAGAGCAACACCACTGGGAC 1747488  
QY 292 TGAACACGCGCCAGACTCTCTACGGGAGGCGACAGTGGGGAATTTGCAATGGGCGCA 351  
DB 1747487 TGACACACGCGCCAGACTCTCTACGGGAGGCGACAGTGGGGAATTTGCGCAATGGAGCA 1747428  
QY 352 AGCCTGATCCAGCAATGCGGGGTGAGTGAAGAACGCCCTTGGTGAAGCTCTTGAG 411  
DB 1747427 AGTGTGAGGAGCAACGCGGGGTGATGAAGAGTTTTCGATGTAATGATCTTTGT 1747368  
QY 412 TCGAAGAAAGAGTTACGTAAATATCTGACTCATGACGTATCGACAGAAAGCA 471  
DB 1747367 TGAAGAAAGAGTTACGTAAATATCTGACTCATGACGTATCGACAGAAAGCA 1747308  
QY 472 CCGGCTTACGTCGTCGACGACCGCGGTAAATGAGGTGTCGCAAGCTTAAATCGGAAT 531  
DB 1747307 ACGGCTTACGTCGTCGACGACCGCGGTAAATGAGGTGTCGCAAGCTTAAATCGGAAT 1747248  
QY 532 ACTGGGCTTAAAGGTGGCGGCGGCTTTTAAAGTCAATGATGAATCCCGGCGCTTA 591  
DB 1747247 ATTGGGCTTAAAGGTGGCGGCGGCTTTTAAAGTCAATGATGAATCCCGGCTTA 1747188  
QY 592 CCGGGAATTCGTTTGAATCTACAGAGCTAGAGTGGCAGAGGAGAGTGAATTCAT 651  
DB 1747187 CCGGGAATTCGTTTGAATCTACAGAGCTAGAGTGGCAGAGGAGAGTGAATTCAT 1747128  
QY 652 GTGTAGCAGTAATGCTGTAAGATATGAAGAACATTCATGTCGAGAGGCACTCTCGG 711  
DB 1747127 GTGTAGCAGTAATGCTGTAAGATATGAAGAACATTCATGTCGAGAGGCACTCTCGG 1747068  
QY 712 GTTAACAATGACGCTGACAGCAAGAAAGGTGGGAGCAAGAGTTGATTAACCTGCTA 771  
DB 1747067 TCTGTAACTGACGCTGACAGCAAGAAAGGTGGGAGCAAGAGTTGATTAACCTGCTA 1747008  
QY 772 GTTCAAGCTTAAAGATGTCATGTAATGTTT--GGGCTTATTAAGCTTGTGAAG 829  
DB 1747007 GTTCAAGCTTAAAGATGTCATGTAATGTTT--GGGCTTATTAAGCTTGTGAAG 1746948  
QY 830 CTAACGCTGAAGTTGACGCGCTGGGAGTACGCTGCGCAATTTAAACTCAAGGAAT 889  
DB 1746947 CTAACGCTGAAGTTGACGCGCTGGGAGTACGCTGCGCAATTTAAACTCAAGGAAT 1746888  
QY 890 GACGGGACCGGCAACAGCGGTGATTTATGAGTTAATTTGAGTCAACGGGAAACCT 949  
DB 1746887 GACGGGACCGGCAACAGCGGTGATTTATGAGTTAATTTGAGTCAACGGGAAACCT 1746828  
QY 950 TACTTAACCTTGAATGATGAGCAATTTTCTAGAGATAGA-TTAGTGCTTCGGGAACGCTA 1008

Db	1746827	TACAGCGTCTTGACAATCCCTTTGACCACCTGTGGAAACAGAGCTTTCCTTGGGGACCAAG	1746768
OY	1009	ACACAGGTCGTGCATGCGCTGTGCTGCAGCTCCTGTGTAAGATGTTGGGTTAAGTCCCGCA	1068
Db	1746767	TGACAGGTGGTGCATGGTGTGTGTCAGACTCGTGTGTAAGATGTTGGGTTAAGTCCCGCA	1746708
OY	1069	ACGAGCGCAACCCCTTGTTCATTAAATTGCCATCATTTTGGTGGGCACTTTAATGACACTGCC	1128
Db	1746707	ACGAGCGCAACCCCTGTGATTATTAAGTGTCCAGCATTTTAAGTGGGCACTCTAAAGTACTGCC	1746488
OY	1129	GSTACAAACCGGAGGAAGTGGGGGATGACGTCACAGTCTCATATGAGCCCTTAATGGGTAGG	1188
Db	1746647	GGTG-CAAGCCGGAGAGAGGTGGGGATACGTCCTCAATTCATCATGCCCCCTTAATGACTGGG	1746589
OY	1189	CTTCACACGTATAACATGCGCGGTACAGAGGGTTGCCAACCCGCGAGGGGAGACTAATC	1248
Db	1746588	CTACACACGTGCTCAATATGATAGTAAACAAAGGTCGCGAACCCGAGAGTGGACTAATC	1746529
OY	1249	TCAGAAAGCGCGTCTGTAGTCCGGATCCGAGTCTGCAAATTGCATCCGTAAGTGGGAATC	1308
Db	1746528	CCATTAACCTATTCTCACGTGTGGGATTTGAGGCTGCACATCTGCCTACATGAAGCCGGAAATC	1746469
OY	1309	GCTAGTAATCGCGGATTCAGCAATGTGCGCGTGAATTAGTTCGCCGGGTCTTGTACACACGCG	1368
Db	1746468	GCTAGTAATCGTGGATTCAGCAATGCCACGCTGAATATCGTTCCCGGGCTTGTACACACCGCG	1746409
OY	1369	CCGTACACCAATGGGAGTGGGTTTTACACAGAAGCAGTAGTCTTAAC- GTAAGAGGAGCG	1427
Db	1746408	CCGTACACCAACGAGAGTTTGTAAACACCGAAGTCGTAAGGTTAACTTTATGAGGCCAG	1746349
OY	1428	CTTGCCACCGGTGAGTATCATGACTGGGGTG	1457
Db	1746348	CCGCCGAAGTGGGACAGATTAATTTGGGGTG	1746319

```

US-11-151-847-1
; Sequence 1, Application US/11151847
; Publication No. US20060002909A1
; GENERAL INFORMATION:
; APPLICANT: Kabushiki Kaisha Yonezawa Birtu Shituteemu Sabisu
; TITLE OF INVENTION: A microorganism separated from kefir grains, a microorganism cult
; TITLE OF INVENTION: obtained by culturing this microorganism or microorganisms inclu
; TITLE OF INVENTION: and a product using them.
; FILE REFERENCE: H16P1032
; CURRENT APPLICATION NUMBER: US/11/151,847
; CURRENT FILING DATE: 2005-06-13
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 1521
; TYPE: DNA
; ORGANISM: Lactobacillus mali
US-11-151-847-1

```

Query Match	56.1%;	Score 817.2;	DB 17;	Length 1521;
Best Local Similarity	76.2%;	Pred. No. 4.1e-245;		
Matches 1069;	Conservative	1;	Mismatches 319;	Indels 14;
				Gaps 5

[illegible]

Db	280	ACCAAGCAGTGAATCCTAGCCCACTGAGAGGTTGATCCGCCCACTTTGGGACCTGAGACA	339
Qy	299	CGGCCGACAGCTCTTACCGGAGGCGACAGTGGGGATTTTGGACATATGCGCAAGCTTGA	358
Db	340	CGGCCCAAACTCTTACCGGAGGCGACAGTAGGGAATCTTCCACATAGGACGCAAGTCTGA	399
Qy	359	TCACGAATGCCGGTGAAGTGAAGAAGGCCCTTCGGGTGTAAACTCTTTCAGTCGAGAA	418
Db	400	TGAGCAACCGCCGGGTGATGAMAAAGTTTTCGGAATCTGTAACTCTGTGTATGAGAA	459
Qy	419	GAAGAAGTTACGGTAATTAATCGTAGCTCATGACGGTATCGACAGAAAGAACCGGCTA	478
Db	460	GAACCTGTGTAGAGTAATCTTCTTACATGATGACGGTATCTTACCAAGAAACACCGGCTA	519
Qy	479	ACTACGTCCAGCAGACCGCCGGTAAATCGTAGGGTGCACAGCCTTAATCGGAATTA	538
Db	520	ACTAGTCCAGCAGACCGCCGGTAAATCGTAGGGTGCACAGCCTTGTCCGGAATTAATGGGC	579
Qy	539	GTAAAGGTGGCGCAGGCGGCTTTTGAATGATGATGAAATCCCGGGCTTAACTGGGA	598
Db	580	GTAAAGGGAACGACGCGCTCTTTTAATCTGAATGTAAACCTTCGGCTTAACTGGAATG	639
Qy	599	AATGGCTTTGAAACTAACAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTATGC	658
Db	640	CGTGATTTGAAACTTGGAGACTTGAAGTGCAGAAAGAGAGTGAATCTCATGTATGC	699
Qy	659	AGTGAATATGCGTAGATATGAGAAACATCATGCGAAGCGACGCTCTGGGTTAA	718
Db	700	GGTGAATATGCGTAGATATATGGAAGAACACCATGTGGCAAGCGGCTCTGTGCTGTAA	759
Qy	719	CTGACGCTCATGACGAAAGCGTGGGGAGCAACAGATTAAGTATCCCTGTATGTCACG	778
Db	760	CTGACGCTGAGGTTTGAAGAGGTGGGTGCAACAGATTAGATACCTGTGTATGTCACG	819
Qy	779	CCCTAAAGATGTCAACTGATTTGTGGCCCTTATTAAGC--TTGGTAAAGAACTTAAACG	836
Db	820	CTGTAAAGATTAATGCTTAAGTGTGTGAGGGTTTCCGCCCTTCAAGTCTGACGCTAAACG	879
Qy	837	GTGAAGTTGACCGCCTGGGGAGTACGATCGCAAGTTAAATCTCAAGAAATTTGACGGG	896
Db	880	ATTAAAGCATCCGCTCGGGAGTACGACCGCAAGGTTAAATCTCAAGAAATTTGACGGG	939
Qy	897	ACCGCACAGCGGTGATTTATGTGGAATTAATCGATGCAAGCGCAAAACCTTACCTTAC	956
Db	940	GCCCCACAGCGGTGAGCATGTGTTTAATTCAMACCAACGCAAAACCTTACCAAG	999
Qy	957	CCTTGAACATGTAAGCAATTTCTAGAGAT-AGATATGCTTCGGGAACGCTAACACAG	1015
Db	1000	CTTTGACATCTTTTCTTAACTGAAAGATCAGGATTTCCCTTCGGGGACAAATATGACAG	105
Qy	1016	TGCTGCATGGCTGTGTCAGCTCGTGTGCTGATGTTGGTTAAAGTCCCGCAAGACG	1075
Db	1060	TGCTGATGTGTGTGTCAGCTCGTGTGCTGATGTTGGTTAAAGTCCCGCAAGACG	1111
Qy	1076	CAACCTGTCTAATTAATGTCATCTTGTGTTGGGACTTTAATGAGACTGCGCGGTGAC	1133
Db	1120	CAACCTTAATTAATGTTGCCAGCACTTAATGTTGGGCACTTAATGAGACTGCGCGGTGAC	1177
Qy	1136	AACCGGAGGAGGTGGGATGACGTCAGTCTCATGAGCCCTTAATGAGGTGAGGCTTACAC	1199
Db	1180	AACCGGAGGAGGTGGGATGACGTCAGTCTCATGAGCCCTTAATGAGGTGAGGCTTACAC	1231
Qy	1196	CGTAATACATGAGCGGTGACAGAGGGTTCCAACCGCGAGGGGAGACTAATCTCAGAA	1251
Db	1240	CGTGTCTAATGAGCGGTGACAGAGGTTCGCAACCGCGAGGGTTCGTAATCTCTTAA	1299
Qy	1256	GGCGCTGTATGTCGGATGGAATCTGCACTCGTGAATGCGGAATCGCTATGTA	1311
Db	1300	ACCGTTCTAGTTCCGATTTGTAGGCTGCAACTCGCTAATGATCGGAATCGCTATGTA	1351
Qy	1316	ATGCGGAGTACAGATGCGGGTGAATAGGTTCCGGGGCTTGTACACCGCCGCTAC	1375
Db	1360	ATGCGGAGTACAGATGCGGGTGAATAGGTTCCGGGGCTTGTACACCGCCGCTAC	1411



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 5, 2006, 13:14:37 ; Search time 919 Seconds  
(without alignments)  
10566.330 Million cell updates/sec

Title: US-10-659-980A-1  
Sequence: 1 atgagcgtggtggcgtatgc.....tgaatctacgtcgtggtg 1457

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues  
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 70 summaries

Database : N\_Geneseq\_21:\*

- 1: geneseqn1980a:\*
- 2: geneseqn1990a:\*
- 3: geneseqn2000a:\*
- 4: geneseqn2001a:\*
- 5: geneseqn2001b:\*
- 6: geneseqn2002a:\*
- 7: geneseqn2002b:\*
- 8: geneseqn2003a:\*
- 9: geneseqn2003b:\*
- 10: geneseqn2003c:\*
- 11: geneseqn2003d:\*
- 12: geneseqn2004a:\*
- 13: geneseqn2004b:\*
- 14: geneseqn2005a:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1457	100.0	1457	6	ABA02416
2	1457	100.0	1457	12	ADM32704
3	1452.2	99.7	1457	6	ABA02417
4	1452.2	99.7	1457	12	ADM32705
5	1391.6	95.5	1494	12	ADM32723
6	1359.2	93.3	1494	12	ADM32722
7	1358.8	93.3	1467	12	ADM32721
8	1287.6	88.4	1458	6	ABA02418
9	1287.6	88.4	1458	12	ADM32706
10	1230.2	84.4	1460	6	ABA02419
11	1230.2	84.4	1460	12	ADM32707
12	1206.4	82.8	1485	4	AAC86030
13	1197.2	82.2	1495	6	ABQ78660
14	1194	81.9	1460	4	AAC86026
15	1190.8	81.7	1464	6	ABL40355
16	1181	80.9	1459	4	AAC86028
17	1179.2	80.9	1460	4	AAC86024
18	1161	79.7	1530	13	ADR45500
19	1156.8	79.4	1460	4	AAC86022

# ALIGNMENTS

RESULT 1	ABA02416	standard; DNA; 1457 BP.
XX	ABA02416	
AC	ABA02416	
XX	29-AUG-2003 (revised)	
DT	04-MAR-2002 (first entry)	
XX	Type A ammonia-oxidising bacterium 16S rRNA gene sequence, R7clone140.	
DE	Type A: ammonia-oxidising bacterium; AOB, nitrite; 16S rRNA gene;	
XX	ribosomal RNA; aquarium; wastewater treatment;	
KM	bioremediation; ds.	
XX		
OS	Nitrosomonadales.	
XX		

20	1155.8	79.3	1460	4	AAC86021	Aac86021 R. tenuis
21	1152.4	79.1	1532	2	AAQ26729	Aaq26729 16S rRNA
22	1152.4	79.1	1532	2	ADB61686	Adb61686 16S rRNA
23	1148.8	78.8	1526	2	AAU01866	Aau01866 P. cepacia
24	1143.8	78.5	1496	12	ADQ16355	Adq16355 Nucleotid
25	1143.8	78.5	1496	14	AEA01071	Aea01071 Bordetell
26	1142	78.4	1426	2	AAC86029	Aac86029 16S rRNA
27	1141.2	78.3	1426	2	AAQ64008	Aaq64008 16S rRNA
28	1138.4	78.1	1535	2	AAV24295	Aav24295 Burkholde
29	1138.4	78.1	1535	2	ADB61689	Adb61689 16S rRNA
30	1137.6	78.1	1530	14	ADM12666	Adm12666 Variovora
31	1137.2	78.1	1485	10	ADB61687	Adb61687 16S rRNA
32	1133	77.8	1400	8	ABZ69299	Abz69299 J. lividum
33	1131.4	77.7	1464	4	ADB61688	Adb61688 16S rRNA
34	1130.6	77.6	1478	4	AAC86023	Aac86023 R. purpur
35	1129.8	77.5	1400	8	ABZ69300	Abz69300 J. lividum
36	1129.2	77.5	1481	10	ADE11077	Adel1077 Acidovora
37	1128.2	77.4	1400	8	ABZ69298	Abz69298 J. lividum
38	1127.6	77.4	1509	12	ADO61789	Ado61789 Acidovora
39	1126.8	77.3	1488	10	ADB61690	Adb61690 16S rRNA
40	1126	77.3	1509	12	ADQ61789	Adq61789 Acidovora
41	1124.6	77.2	1463	10	ADC53929	Adc53929 Phenyl hy
42	1124.6	77.2	1482	8	ABZ69304	Abz69304 Jantlinob
43	1124.6	77.2	1544	10	ADB61693	Adb61693 16S rRNA
44	1124.6	77.2	110000	3	AAA81490_00	AAA81490 N. mening
45	1124.6	77.2	110000	3	AAA81490_02	AAA81490 (3 of
46	1124.6	77.2	110000	3	AAA81490_03	AAA81490 (4 of
47	1124.6	77.2	110000	3	AAA81489_2	AAA81489 (3 of
48	1124.6	77.2	110000	3	AAA81489_6	AAA81489 (7 of
49	1124.6	77.2	110000	3	AAA81489_7	AAA81489 (8 of
50	1124.6	77.2	172325	3	AAE21613	Aae21613 Neisseria
51	1124.6	77.2	349980	3	AAE21617	Aae21617 Neisseria
52	1124.6	77.2	349980	3	AAE21611	Aae21611 Neisseria
53	1124.6	77.2	349980	3	AAE21612	Aae21612 Neisseria
54	1124.6	77.2	349980	3	AAE21544	Aae21544 Neisseria
55	1123.2	77.1	1400	6	ABV72366	Abv72366 J. lividum
56	1123	77.1	1452	8	ABV72366	Abv72366 Nucleotid
57	1122	77.1	1610	10	ADB61691	Adb61691 16S rRNA
58	1119.8	76.9	1544	5	AA811025	AA811025 Neisseria
59	1119.8	76.9	1544	10	ADB61692	Adb61692 16S rRNA
60	1119.2	76.8	1477	2	AAK26285	Aak26285 Pseudom
61	1118.2	76.7	1536	2	AAI18765	Aai18765 Pseudom
62	1108	76.0	1451	4	AAE74542	Aae74542 Burkholde
63	1108	76.0	1451	10	ABX10819	Abx10819 Burkholde
64	1108	76.0	1451	12	ADH74483	Adh74483 TPA produ
65	1104.4	75.8	1453	6	ABU53377	Abu53377 Escherich
66	1103.2	75.7	1453	10	ADG44143	Adg44143 Unknown o
67	1103.2	75.7	1453	10	ADG17998	Adg17998 Unknown b
68	1103.2	75.7	1453	11	ADL27933	Adl27933 RAZ 16S r
69	1103.2	75.7	1453	12	ADL47789	Adl47789 Unknown b
70	1103	75.7	1440	14	ADL14078	Adl14078 Chromobac

Query Match	100.0%	Score 1457	DB 6	Length 1457
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1457	Conservative 0	Mismatches 0	Indels 0	Gaps 0
DB	1	ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTGGACGGCAGCAGCGATGCTTGCA	60	
DB	1	ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTGGACGGCAGCAGCGATGCTTGCA	60	
DB	61	CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCATCGGAACTATCCAGAAAGGGGGGTA	120	
DB	61	CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCATCGGAACTATCCAGAAAGGGGGGTA	120	
DB	121	ACGCATCGAAGATGCTCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAA	180	
DB	121	ACGCATCGAAGATGCTCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAA	180	
DB	181	GACCTTGGGCTTTTGAAGCGCGCCGATGCTGATTAAGTGAAGTGGGGTAAAGGCTTAC	240	
DB	181	GACCTTGGGCTTTTGAAGCGCGCCGATGCTGATTAAGTGAAGTGGGGTAAAGGCTTAC	240	
DB	241	CAAGCGACGATCAGTATGTTGGTCTGAGAGAGACGACGACGACCTGAGGACTGAGACACG	300	
DB	241	CAAGCGACGATCAGTATGTTGGTCTGAGAGAGACGACGACGACCTGAGGACTGAGACACG	300	
DB	301	GCCGAGACTCTTAACGGGAGGACGAGTGGGGAATTTTGGACAAATGGGGCCAGGCTGATC	360	
DB	301	GCCGAGACTCTTAACGGGAGGACGAGTGGGGAATTTTGGACAAATGGGGCCAGGCTGATC	360	
DB	361	CAGCAATGCGCGGTGAGTGAAGAAAGGCTTGGGTTGTAAAGCTCTTTCACTGAGAGAA	420	
DB	361	CAGCAATGCGCGGTGAGTGAAGAAAGGCTTGGGTTGTAAAGCTCTTTCACTGAGAGAA	420	
DB	421	AAAGGTACCGGTAAATTAATCGTGAATCATGACGGTATCGACAGAAAGAACACCGGCTAAC	480	

Db	421	AAAGGTTACGGTAAATTAATCGTACTCTAGACGGTATGACAGAAAGAACCGGCTAAC	480
Qy	481	TACGTCCAGACAGCCCGCGTAAATACGTAGGGGTGCAGAGCGTTAATCGAATTACTG3GCGT	540
Db	481	TACGTGCCACAGACCGCGCGTAAATACGTAGGGGTGCAGAGCGTTAATCGAATTACTG3GCGT	540
Qy	541	AAAGGTCGCCAGAGCGGCGCTTTGTAAAGTACAGATGTGAAAATCCCGGGGCTTAACCTGTGGAAAT	600
Db	541	AAAGGTCGCCAGAGCGGCGCTTTGTAAAGTACAGATGTGAAAATCCCGGGGCTTAACCTGTGGAAAT	600
Qy	601	TGCGTTTGAACCTACCAAGGCTAGAGTGTGGACAGAGGAGTGGAATTCATCTGTATGACAG	660
Db	601	TGCGTTTGAACCTACCAAGGCTAGAGTGTGGACAGAGGAGTGGAATTCATCTGTATGACAG	660
Qy	661	TGAATTCGGTAAAGATATATGAAAGAACATCGATGGCGAAGGCGACGCTCTCGGGTTAACT	720
Db	661	TGAATTCGGTAAAGATATATGAAAGAACATCGATGGCGAAGGCGACGCTCTCGGGTTAACT	720
Qy	721	GACGCTCATGCAAGAACCGTGGGGAGCAAAACAGATTAATAGTAACTCCCTGTATCCACGCGC	780
Db	721	GACGCTCATGCAAGAACCGTGGGGAGCAAAACAGATTAATAGTAACTCCCTGTATCCACGCGC	780
Qy	781	CTAAACGATGTCAACTAGTGTGTGGGCGCTTATTAGGCTTGGTAAAGAACTAACCGGTGA	840
Db	781	CTAAACGATGTCAACTAGTGTGTGGGCGCTTATTAGGCTTGGTAAAGAACTAACCGGTGA	840
Qy	841	AGTTGACCGGCTGGGGAGTACGGTGCAGAAATTTAAACTCAAAGGAATTGACGGGGACCC	900
Db	841	AGTTGACCGGCTGGGGAGTACGGTGCAGAAATTTAAACTCAAAGGAATTGACGGGGACCC	900
Qy	901	GCAACAAGGGGTGTGATTAATGTGGAATTAATTCGATGCAAGCGCAAAAACCTTACCTAACCTT	960
Db	901	GCAACAAGGGGTGTGATTAATGTGGAATTAATTCGATGCAAGCGCAAAAACCTTACCTAACCTT	960
Qy	961	GACATGTAGCAAAATTTTCTAGAGATAGATTAGTGTCTGGGAGACGCTTAACACAGGTGCTG	1020
Db	961	GACATGTAGCAAAATTTTCTAGAGATAGATTAGTGTCTGGGAGACGCTTAACACAGGTGCTG	1020
Qy	1021	CATGGCTGTGCTCAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACC	1080
Db	1021	CATGGCTGTGCTCAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACC	1080
Qy	1081	CTTGTGCATTAATTTGGCATCATTTTGGTGTGGGCACTTTAATGAGACTGCGGTGACAAACCG	1144
Db	1081	CTTGTGCATTAATTTGGCATCATTTTGGTGTGGGCACTTTAATGAGACTGCGGTGACAAACCG	1144
Qy	1141	GAGGAAGGTGGGGAGTACGTCAAGTCCCTCATGSCCTTATGGGGTGTGGGCTTACACGTTA	1200
Db	1141	GAGGAAGGTGGGGAGTACGTCAAGTCCCTCATGSCCTTATGGGGTGTGGGCTTACACGTTA	1200
Qy	1201	TACAATGCGCGGTACAGAGGGTTCGCCAACCCGCGAGGGGGAGCTAATCTCAGAAAGCGCG	1266
Db	1201	TACAATGCGCGGTACAGAGGGTTCGCCAACCCGCGAGGGGGAGCTAATCTCAGAAAGCGCG	1266
Qy	1261	TCGTATGTCGGGATCGGAATCTTGCAACTGCACTCCGTGAAAGTTCGGAATGCGTAGTAATGCG	1322
Db	1261	TCGTATGTCGGGATCGGAATCTTGCAACTGCACTCCGTGAAAGTTCGGAATGCGTAGTAATGCG	1322
Qy	1321	GGATCAGCATGTGCGGGGTGAATACGTTCCGGGGTCTTGTACACACCGCCCGTCAACCAAT	1388
Db	1321	GGATCAGCATGTGCGGGGTGAATACGTTCCGGGGTCTTGTACACACCGCCCGTCAACCAAT	1388
Qy	1381	GGGAGTGGGTTTCAACCAAGACAGGTAGTCTTAACGTTAAGGAGGGCGCTTGCCACGGTGA	1444
Db	1381	GGGAGTGGGTTTCAACCAAGACAGGTAGTCTTAACGTTAAGGAGGGCGCTTGCCACGGTGA	1444
Qy	1441	GATTTCATGACTGGGGGTG 1457	
Db	1441	GATTTCATGACTGGGGGTG 1457	

## RESULT 2

ADM32704  
ID ADM32704 standard; DNA; 1457 BP.  
XX  
AC ADM32704;  
XX  
XX 17-JUN-2004 (first entry)  
XX  
DE AOB Type A R7clone140 16S rDNA.  
XX  
XX 16S rDNA; ammonia oxidizing bacteria; AOB; ammonia; nitrite;  
XX aqueous environment; freshwater; seawater; aquarium; ss.  
XX  
XX Nitrosomonas sp.  
XX OS  
XX PN MO2004026772-AZ.  
XX PD 01-APR-2004.  
XX PF 10-SEP-2003; 2003WO-US028210.  
XX PR 19-SEP-2002; 2002US-0386217P.  
XX PR 19-SEP-2002; 2002US-0386218P.  
XX PR 19-SEP-2002; 2002US-0386219P.  
XX PA (AQUA-) AQUARIA INC.  
XX  
PI Hovanec TA;  
XX  
DR WPI; 2004-304936/28.  
XX  
PT New composition comprising an isolated bacterial strain that oxidizes  
PT ammonia to nitrite, useful for alleviating or preventing the accumulation  
PT of ammonia in aqueous environment.  
XX  
XX Disclosure; Page 8-9; 98pp; English.  
XX  
XX This sequence represents a 16S rDNA sequence derived from an ammonia  
XX oxidizing bacteria (AOB). This sequence may be used in a composition  
XX which comprises an isolated bacterial strain that oxidizes ammonia to  
XX nitrite. The composition may be used for alleviating or preventing the  
XX accumulation of ammonia in a medium. The ammonia is reduced by at least  
XX 30% when compared with a level of ammonia that would exist in the absence  
XX of the bacterial strain. The composition is useful for alleviating or  
XX preventing the accumulation of ammonia in aqueous environment, e.g. a  
XX freshwater or seawater aquarium.  
XX  
SQ Sequence 1457 BP; 378 A; 315 C; 456 G; 308 T; 0 U; 0 Other:  
Query Match 100.0%; Score 1457; DB 12; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATTGAACGCTGGCGGATGCTTTACATGCAAGTCGAACGGCAGCGATGCTTCAT 60  
DB 1 ATTGAACGCTGGCGGATGCTTTACATGCAAGTCGAACGGCAGCGATGCTTCAT 60  
QY 61 CTGGTGGCGAGTGGCGGACGGGTGAATGATCGAAGCGTATCCGAAGAGGGGGGTA 120  
DB 61 CTGGTGGCGAGTGGCGGACGGGTGAATGATCGAAGCGTATCCGAAGAGGGGGGTA 120  
QY 121 ACCGATGAAAGATGCTCTAATACCGCATATCTTAAGAGGAAAGCAGGGGATCGAAA 180  
DB 121 ACCGATGAAAGATGCTCTAATACCGCATATCTTAAGAGGAAAGCAGGGGATCGAAA 180  
QY 181 GACCTTCGCTTTTGAAGCGGCGCATGTCTGATTAAGCTAATGGTGGGCTTAAGCCTAC 240  
DB 181 GACCTTCGCTTTTGAAGCGGCGCATGTCTGATTAAGCTAATGGTGGGCTTAAGCCTAC 240  
QY 241 CAAGCGCAGATCACTGTTGGTCTGAAGGAGCAGCAGCCACTGGGAGCTGAGACAGC 300  
DB 241 CAAGCGCAGATCACTGTTGGTCTGAAGGAGCAGCAGCCACTGGGAGCTGAGACAGC 300  
QY 301 GCCCAGACTCTTACGGGAGCAGCAGTGGGAAATTTTGAACATGGGCGCAAGCCTGATC 360

DB 301 GCCCAGACTCTTACGGGAGCAGCAGTGGGAAATTTTGAACATGGGCGCAAGCCTGATC 360  
QY 361 CAGCAATCCCGGTGAGTGAAGAAAGGCTTCGGGTGTTAAAGCTTTTCACTCGAAGAA 420  
DB 361 CAGCAATCCCGGTGAGTGAAGAAAGGCTTCGGGTGTTAAAGCTTTTCACTCGAAGAA 420  
QY 421 AAAGTTACGGTAAATTAATCGTACTCATGACGGTATGACAGAAAGACCGGCTAAC 480  
DB 421 AAAGTTACGGTAAATTAATCGTACTCATGACGGTATGACAGAAAGACCGGCTAAC 480  
QY 481 TACGTCACAGCAGCCCGGTAAATACGTAGGGTCAAGGCTTAATCGAATTACTGGCGT 540  
DB 481 TACGTCACAGCAGCCCGGTAAATACGTAGGGTCAAGGCTTAATCGAATTACTGGCGT 540  
QY 541 AAAGGTGGCAGGCGGCTTTGTAAGTCAAGATGTGAAATCCCGGGCTTAACCTGGGAT 600  
DB 541 AAAGGTGGCAGGCGGCTTTGTAAGTCAAGATGTGAAATCCCGGGCTTAACCTGGGAT 600  
QY 601 TCGGTTTGAACCTAACAAGCTTGAAGTGGGAGAGAGGAGTGGAAATTCATGTTAGCAG 660  
DB 601 TCGGTTTGAACCTAACAAGCTTGAAGTGGGAGAGAGTGGAAATTCATGTTAGCAG 660  
QY 661 TGAATTCGCTAGATATGAGAAACATCATGAGCGAAGGACGCTCTCGGTTAACTACT 720  
DB 661 TGAATTCGCTAGATATGAGAAACATCATGAGCGAAGGACGCTCTCGGTTAACTACT 720  
QY 721 GACGCTCATGACGAAGAGCGTGGGAGCAAGAGATTAGATACCTGTTAGTCCAGCC 780  
DB 721 GACGCTCATGACGAAGAGCGTGGGAGCAAGAGATTAGATACCTGTTAGTCCAGCC 780  
QY 781 CTAAACGATGCTAAGTGTGTTGGGCTTTAATTAAGCTTGTATACGAACTAACGGGTGA 840  
DB 781 CTAAACGATGCTAAGTGTGTTGGGCTTTAATTAAGCTTGTATACGAACTAACGGGTGA 840  
QY 841 AGTTGACCGCCCTGGGAGATAGCGTGGCAAGATTAAATCTCAAGAAATGACGGGAGCC 900  
DB 841 AGTTGACCGCCCTGGGAGATAGCGTGGCAAGATTAAATCTCAAGAAATGACGGGAGCC 900  
QY 901 GCACAAAGCGGTGATTAATGATGATTAATTCGATGCAACCGCAAAACCTTACTCACTT 960  
DB 901 GCACAAAGCGGTGATTAATGATGATTAATTCGATGCAACCGCAAAACCTTACTCACTT 960  
QY 961 GACATGATGCAATTTTCTAGAGATGATTAATGATGATGATGATGATGATGATGATGAT 1020  
DB 961 GACATGATGCAATTTTCTAGAGATGATTAATGATGATGATGATGATGATGATGATGAT 1020  
QY 1021 CATGCTGCTGCTAGCTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
DB 1021 CATGCTGCTGCTAGCTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
QY 1081 CTGTGATTAATTCATCATATTTGTTGGGCACTTTAATGAGCTGCGGTGACAAACCG 1140  
DB 1081 CTGTGATTAATTCATCATATTTGTTGGGCACTTTAATGAGCTGCGGTGACAAACCG 1140  
QY 1141 GAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
DB 1141 GAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
QY 1201 TACAAATGGCGGTCAGAGAGGTTGCAACCGGAGGAGGAGTAAATCTCAGAAAGCGC 1260  
DB 1201 TACAAATGGCGGTCAGAGAGGTTGCAACCGGAGGAGGAGTAAATCTCAGAAAGCGC 1260  
QY 1261 TCGTATCGCGATCGAGATCTGCAATCGATCCGTAAGTCCGTAATTCGCTAGTAATTCG 1320  
DB 1261 TCGTATCGCGATCGAGATCTGCAATCGATCCGTAAGTCCGTAATTCGCTAGTAATTCG 1320  
QY 1321 GATCAGCATGTCGCGGTGAATTCGTTCCGGGTCTTTGTAACACACCGCCGCTACACAT 1380  
DB 1321 GATCAGCATGTCGCGGTGAATTCGTTCCGGGTCTTTGTAACACACCGCCGCTACACAT 1380  
QY 1381 GGGAGTGGTTTACCAAGAAAGCATAGCTTAACCGTAAAGGAGGGGCTTGCCACGGTGA 1440



DB 1381 GGGAGTGGCTTTCACACAGACGAGTACTTAACCGTAAAGAGGGCGCTTGCCACGCTGCA 1440

OY 1441 GATTCATGACTGGGGTG 1457  
 |||||  
 |||||

DB 1441 GATTCATGACTGGGGTG 1457

RESULT 3  
 ABA02417  
 ID ABA02417 standard; DNA; 1457 BP.  
 XX ABA02417;  
 AC ABA02417;  
 XX  
 DT 29-AUG-2003 (revised)  
 DT 04-MAR-2002 (first entry)  
 DE Type A1 ammonia-oxidising bacterium 16S rRNA gene sequence, R7clone187.  
 XX  
 KW Type A1; ammonia-oxidising bacterium; AOB; nitrite; 16S rRNA gene;  
 KW ribosomal RNA; aquaria; aquaculture; waste water treatment;  
 KW bioremediation; de.  
 XX  
 OS Nitrosomonadales.  
 XX  
 PN WO200190312-A1.  
 PD 29-NOV-2001.  
 PD 17-MAY-2001; 2001MO-US016265.  
 PF 19-MAY-2000; 2000US-00573684.  
 PR 19-MAY-2000; 2000US-00573684.  
 XX  
 PA (AQUA-) AQUARIA INC.  
 XX  
 PI Hovanec TA, Burrell PC;  
 DR WPI; 2002-075367/10.  
 PT New bacteria capable of oxidizing ammonia to nitrite, for preventing or  
 PT alleviating the accumulation of ammonia in fresh water aquaria, seawater  
 PT aquaria and waste water.  
 XX  
 PS Claim 2; Page 6; 62pp; English.  
 CC The invention relates to 4 novel types of ammonia-oxidising bacteria  
 CC (AOB) found in freshwater aquaria. The bacteria are able to oxidise  
 CC ammonia to nitrite and are members of the ammonia-oxidising bacteria  
 CC family of the beta subdivision of Proteobacteria. The 4 types of bacteria  
 CC can be distinguished on the basis of their 16S rRNA (ribosomal RNA) gene  
 CC sequences (ABA02416-ABA02419), and are classified as AOB type A (e.g.,  
 CC R7clone140), type A1 (e.g., R7clone187), type B (e.g., R3clone5) and type  
 CC C (e.g., R3clone47). The invention also encompasses isolated 16S rRNA  
 CC gene sequences of the ammonia-oxidising bacteria of the invention,  
 CC oligonucleotide probes and primers for the detection of these bacteria,  
 CC and compositions comprising the bacteria. The bacteria of the invention  
 CC are useful in biological filters for reducing ammonia accumulation in  
 CC both freshwater and seawater aquaria. They may also be used in waste  
 CC water treatment and in bioremediation processes to reduce the level of  
 CC pollution caused by ammonia. The present sequence represents R7clone187,  
 CC a 16S rRNA gene sequence from the type A1 ammonia-oxidising bacterium of  
 CC the invention. (Updated on 29-AUG-2003 to standardise OS field)  
 XX  
 SO Sequence 1457 BP; 379 A; 317 C; 455 G; 306 T; 0 U; 0 Other;

Query Match 99.7%; Score 1452.2; DB 6; Length 1457;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0.

OY 1 ATTGAACGGTGGCGGAGCTTTACACATGCAAGTGCAGACGGCAGACGAGATCTTGCAAT 60  
 |||||  
 |||||

DB 1 ATTGAACGGTGGCGGAGCTTTACACATGCAAGTGCAGACGGCAGACGAGATCTTGCAAT 60  
 |||||  
 |||||

OY 61 CTGTGTGGCCAGTGTGGCGAGCGGGTGAAGTATGTGATCGGAAGCTTATCCAGAAAGGAGGAGTA 120

Db	61	CTGGTGGCGAATGGCGGACGGGTGAATATCGAATCGGAACGTATCCAGAAAGAGGGGGGTA	120
Qy	121	ACGCATCGAAAGATGTGCTAATATACCGCATATACTCTAAGAGGAAACGAGGGATCGAA	180
Db	121	ACGCATCGAAAGATGTGCTAATATACCGCATATACTCTAAGAGGAAACGAGGGATCGAA	180
Qy	181	GACCTTGGGCTTTTGGAGCGGGCGGATGTCGATTAAGTAGTAGTTGGTGGGGTAAAGGCTTAC	240
Db	181	GACCTTGGGCTTTTGGAGCGGGCGGATGTCGATTAAGTAGTAGTTGGTGGGGTAAAGGCTTAC	240
Qy	241	CAAGCGCAGATCAGTAGTGGTCTGAGAGGACCAACGACCACTGGGACTGAGACG	300
Db	241	CAAGCGCAGATCAGTAGTGGTCTGAGAGGACCAACGACCACTGGGACTGAGACG	300
Qy	301	GCCCAACTCTTACGGGAGGCGACAGATGGGGAAATTTTGGACMAATGGGCGCAAGCTTATC	360
Db	301	GCCCAACTCTTACGGGAGGCGACAGATGGGGAAATTTTGGACMAATGGGCGCAAGCTTATC	360
Qy	361	CAGCAATCCGCGTGAAGTGAAGAAAGGCTTGGGTTTAAAGCTCTTTCAGTCGAGAGA	420
Db	361	CAGCAATCCGCGTGAAGTGAAGAAAGGCTTGGGTTTAAAGCTCTTTCAGTCGAGAGA	420
Qy	421	AAAGGTTACGGTAAATATATCTGATCTCATGACGGTATCGACAGAAAGCAGCGGCTTAC	480
Db	421	AAAGGTTACGGTAAATATATCTGATCTCATGACGGTATCGACAGAAAGCAGCGGCTTAC	480
Qy	481	TACGTGCCAGAGCGCGGTAAATACGTAAAGGTGCAACGGTTAATCGGAATTACTGGGCGT	540
Db	481	TACGTGCCAGAGCGCGGTAAATACGTAAAGGTGCAACGGTTAATCGGAATTACTGGGCGT	540
Qy	541	AAAGGCTGCGGAGCGGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAAT	600
Db	541	AAAGGCTGCGGAGCGGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACTTGGGAAT	600
Qy	601	TGCGTTTGAAACTACAGGCTTAAGTGTGGCAGAGGGAGGTGGAATTCCTATGTCTACAG	660
Db	601	TGCGTTTGAAACTACAGGCTTAAGTGTGGCAGAGGGAGGTGGAATTCCTATGTCTACAG	660
Qy	661	TGAATGCGTGAATATGAAAGCAATCGATGCGCAAGGACGCTCTGGGTTAACACT	720
Db	661	TGAATGCGTGAATATGAAAGCAATCGATGCGCAAGGACGCTCTGGGTTAACACT	720
Qy	721	GACGCTACGACGAAAGCGTGGGAGCAACAGGATTAATACCTGGTATGTCACAGCC	780
Db	721	GACGCTACGACGAAAGCGTGGGAGCAACAGGATTAATACCTGGTATGTCACAGCC	780
Qy	781	CTAAACGATGCACTAGTTGTTGGGCTTATTAAGGCTTGGTAAAGAGCTTACGCGTGA	840
Db	781	CTAAACGATGCACTAGTTGTTGGGCTTATTAAGGCTTGGTAAAGAGCTTACGCGTGA	840
Qy	841	AGTTGACCGCTGGGGAGTACGGTTCGCAAGATTAAACTCAAGAGAAATGACGGGGACC	900
Db	841	AGTTGACCGCTGGGGAGTACGGTTCGCAAGATTAAACTCAAGAGAAATGACGGGGACC	900
Qy	901	GCACAAGCGGTGATATATGTGATTAATTCGATGCAACGCAAAAACCTTAACCTT	960
Db	901	GCACAAGCGGTGATATATGTGATTAATTCGATGCAACGCAAAAACCTTAACCTT	960
Qy	961	GACATGTGCAATTTTCTAAGATAGATTAGTCTTCGAGAACGCTTAACAAGGTCTG	1020
Db	961	GACATGTGCAATTTTCTAAGATAGATTAGTCTTCGAGAACGCTTAACAAGGTCTG	1020
Qy	1021	CATGGCTGTGCTCAGCTGTGCTGAGATTTGGGTTAACTCCCGCAACGAGCGCAAC	1080
Db	1021	CATGGCTGTGCTCAGCTGTGCTGAGATTTGGGTTAACTCCCGCAACGAGCGCAAC	1080
Qy	1081	CTTGCAATTAATGCAATTTGGTGGGCACTTAAATGAACTGCGCGGTGACAAACG	1140
Db	1081	CTTGCAATTAATGCAATTTGGTGGGCACTTAAATGAACTGCGCGGTGACAAACG	1140
Qy	1141	GAGGAAGGTGGGGATGACGTCAAGTCTCATGACCTTAATGGGTAGGGCTTACACGTAA	1200
Db	1141	GAGGAAGGTGGGGATGACGTCAAGTCTCATGACCTTAATGGGTAGGGCTTACACGTAA	1200



Db 1141 GAGGAAGTGGGAGTGAAGCTCAATCTCATGCGCCCTTAATGGGTAGGCGCTTACAGCTAA 1200  
Qy 1201 TACAAATGCGCGCTGACAGAGGTTGCCAACCCTGGAGGGGAGCTAAATCTCAGAAAACGCG 1260  
Db 1201 TACAAATGCGCGCTGACAGAGGTTGCCAACCCTGGAGGGGAGCTAAATCTCAGAAAACGCG 1260  
Qy 1261 TCGTAGTCCGAGTCCGAGTCTGCAATCGAATCCGTTGAAGTGGAAATCGCTAGTATCCG 1320  
Db 1261 TCGTAGTCCGAGTCCGAGTCTGCAATCGAATCCGTTGAAGTGGAAATCGCTAGTATCCG 1320  
Qy 1321 GGATCGCATGTCGCGCTGAATACGTTCCCGGCTCTGTATACACGCGCCGTCACACAT 1380  
Db 1321 GGATCGCATGTCGCGCTGAATACGTTCCCGGCTCTGTATACACGCGCCGTCACACAT 1380  
Qy 1381 GGGAGTGGGTTTCCAGCAAGACAGTAGTCTAAACCGTAAAGAGAGGCGCTTGGCCACGCTGA 1440  
Db 1381 GGGAGTGGGTTTCCAGCAAGACAGTAGTCTAAACCGTAAAGAGAGGCGCTTGGCCACGCTGA 1440  
Qy 1441 GATTCAATGACTGGGGTG 1457  
Db 1441 GATTCAATGACTGGGGTG 1457  
RESULT 4  
ADM32705 standard; DNA; 1457 BP.  
ID ADM32705; ADM32705;  
XX ADM32705;  
AC ADM32705;  
XX ADM32705;  
DT 17-JUN-2004 (first entry)  
XX 17-JUN-2004 (first entry)  
XX AOB Type A1 R7clone187 16S rDNA.  
DE 16S rDNA; ammonia oxidizing bacteria; AOB; ammonia; nitrite;  
KM aqueous environment; freshwater; seawater; aquarium; ss.  
XX  
XX Nitrosomonas sp.  
OS  
PN W02004026772-A2.  
PD 01-APR-2004.  
XX 01-APR-2004.  
PF 10-SEP-2003; 2003MO-US028210.  
XX 10-SEP-2003; 2003MO-US028210.  
PR 19-SEP-2002; 2002JUS-0386217P.  
PR 19-SEP-2002; 2002JUS-0386218P.  
PR 19-SEP-2002; 2002JUS-0386219P.  
XX 19-SEP-2002; 2002JUS-0386219P.  
XX (AQUA-) AQUARIA INC.  
PA  
XX  
PI Hovaneec TA;  
XX Hovaneec TA;  
DR WPI; 2004-304936/28.  
XX WPI; 2004-304936/28.  
PT New composition comprising an isolated bacterial strain that oxidizes  
PT ammonia to nitrite, useful for alleviating or preventing the accumulation  
PT of ammonia in aqueous environment.  
XX  
XX  
XX Disclosure; Page 10; 98pp; English.  
XX  
XX This sequence represents a 16S rDNA sequence derived from an ammonia  
XX oxidizing bacteria (AOB). This sequence may be used in a composition  
XX which comprises an isolated bacterial strain that oxidizes ammonia to  
XX nitrite. The composition may be used for alleviating or preventing the  
XX accumulation of ammonia in a medium. The ammonia is reduced by at least  
XX 30% when compared with a level of ammonia that would exist in the absence  
XX of the bacterial strain. The composition is useful for alleviating or  
XX preventing the accumulation of ammonia in aqueous environment, e.g. a  
XX freshwater or seawater aquarium.  
XX  
XX Sequence 1457 BP; 379 A; 317 C; 455 G; 306 T; 0 U; 0 Other;  
XX  
XX Query Match 99.7%; Score 1452.2; DB 12; Length 1457;

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 ATTTGAACCTGGCGGCATGCTTTTACATGCAATGCAAGTCGAACCGGACGACGAGTGTGAT 60  
Db 1 ATTTGAACCTGGCGGCATGCTTTTACATGCAATGCAAGTCGAACCGGACGACGAGTGTGAT 60  
Qy 61 CTGGTGGGAGTGGGGAACGGGTGATATGATCGAAGCGATCCAGAGAGGGGGTGA 120  
Db 61 CTGGTGGGAGTGGGGAACGGGTGATATGATCGAAGCGATCCAGAGAGGGGGTGA 120  
Qy 121 AGCGATCGAAGATGTGCTAATACCGCATATATCTTAAAGAGAAAGCAGGGGATCGAAA 180  
Db 121 AGCGATCGAAGATGTGCTAATACCGCATATATCTTAAAGAGAAAGCAGGGGATCGAAA 180  
Qy 181 GACCTTGGCGCTTTTGAAGCGCGCATGTCTGTATAGCTAGTGTGGGTAAAGGCTTAC 240  
Db 181 GACCTTGGCGCTTTTGAAGCGCGCATGTCTGTATAGCTAGTGTGGGTAAAGGCTTAC 240  
Qy 241 CAAGCGAGATCAGTATGTTGCTGAGAGAGCGACGACCACTGGGACTGAGACAG 300  
Db 241 CAAGCGAGATCAGTATGTTGCTGAGAGAGCGACGACCACTGGGACTGAGACAG 300  
Qy 301 GCCAGACTCTTACGCGGAGGACAGCATGCGGAAATTTTGAACAATGGGCGCAAGCTGATC 360  
Db 301 GCCAGACTCTTACGCGGAGGACAGCATGCGGAAATTTTGAACAATGGGCGCAAGCTGATC 360  
Qy 361 CAGCAATGCGCGGTGATGTAAGAGAGCGCTTGGGTTTAAAGCTCTTTCAATGTCAGAA 420  
Db 361 CAGCAATGCGCGGTGATGTAAGAGAGCGCTTGGGTTTAAAGCTCTTTCAATGTCAGAA 420  
Qy 421 AAAGTTACCGTAAATATCGTGAATCATGACGGTATCGACAGAAAGACACCGGCTAAC 480  
Db 421 AAAGTTACCGTAAATATCGTGAATCATGACGGTATCGACAGAAAGACACCGGCTAAC 480  
Qy 481 AAAGTTACCGTAAATATCGTGAATCATGACGGTATCGACAGAAAGACACCGGCTAAC 480  
Db 481 AAAGTTACCGTAAATATCGTGAATCATGACGGTATCGACAGAAAGACACCGGCTAAC 480  
Qy 481 TACGTCCAGACGCGCGGTAAATACGTAGGTTGCAAGCGTTAATCGAAATTAATGCGCGT 540  
Db 481 TACGTCCAGACGCGCGGTAAATACGTAGGTTGCAAGCGTTAATCGAAATTAATGCGCGT 540  
Qy 541 AAAGGTGCGCAGCGCGCTTTTGAATCAATGTAATATCCCGGCTTAACTGGGAAAT 600  
Db 541 AAAGGTGCGCAGCGCGCTTTTGAATCAATGTAATATCCCGGCTTAACTGGGAAAT 600  
Qy 601 TCGGTTTGAATCAAGCTGATGATGTCGAGAGGAGTGGAAATTCATGTTAGCAG 660  
Db 601 TCGGTTTGAATCAAGCTGATGATGTCGAGAGGAGTGGAAATTCATGTTAGCAG 660  
Qy 661 TGAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 661 TGAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGACGCC 780  
Db 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGACGCC 780  
Qy 781 CTAACGATGCTCACTAGTTGTTGGGCTTATTAAGCTTGTATACCAAGTAAACGGTGA 840  
Db 781 CTAACGATGCTCACTAGTTGTTGGGCTTATTAAGCTTGTATACCAAGTAAACGGTGA 840  
Qy 841 AGTTGACCGCTGGGGAGTACGCTGCGCAAGATTAACCAAGAAATTCAGCGGAGCC 900  
Db 841 AGTTGACCGCTGGGGAGTACGCTGCGCAAGATTAACCAAGAAATTCAGCGGAGCC 900  
Qy 901 GCAACAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAACCTTAACCTACCTT 960  
Db 901 GCAACAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAACCTTAACCTACCTT 960  
Qy 961 GACATGTAGGAAATTTCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db 961 GACATGTAGGAAATTTCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Qy 1021 CATGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
Db 1021 CATGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080

Db 1021 CATGCTGTCGTCAGCTCGTGTGAGATGTTGGTTAAGTCCCGCAACGACGCCAACC 1080  
Qy 1081 CTTGTCATTAAATTCATCATTTGGTTGGGCACTTAAAGACCTGCCGGTACAAACCG 1140  
Db 1081 CTTGTCATTAAATTCATCATTTGGTTGGGCACTTAAAGACCTGCCGGTACAAACCG 1140  
Qy 1141 GAGGAAGTGGGAGATGACGTCAAGTCTCATGCGCCCTTATGGGTAGGGCTTCACAGTAA 1200  
Db 1141 GAGGAAGTGGGAGATGACGTCAAGTCTCATGCGCCCTTATGGGTAGGGCTTCACAGTAA 1200  
Qy 1201 TACAAATGGCGGTACAGAGGGTTGCCAACCCCGAGGGGAGCTAACTTCAGAAAGCGCG 1260  
Db 1201 TACAAATGGCGGTACAGAGGGTTGCCAACCCCGAGGGGAGCTAACTTCAGAAAGCGCG 1260  
Qy 1261 TCGTAGTCCGGATCCGAAGTCTGCAATCTGCACTCCGTAAGTGGGAATGGCTATGTAATGCC 1320  
Db 1261 TCGTAGTCCGGATCCGAAGTCTGCAATCTGCACTCCGTAAGTGGGAATGGCTATGTAATGCC 1320  
Qy 1321 GGATCAGCATGTGCGCGGTAAATACGTTCCGGGTCTTTTACACACCGCCGTCACACCAT 1380  
Db 1321 GGATCAGCATGTGCGCGGTAAATACGTTCCGGGTCTTTTACACACCGCCGTCACACCAT 1380  
Qy 1381 GGGAGTGGGTTTACCAAGAGCAGGTAGTCTTAACCGTAAGAGAGGGCGCTTGCACGGTGA 1440  
Db 1381 GGGAGTGGGTTTACCAAGAGCAGGTAGTCTTAACCGTAAGAGAGGGCGCTTGCACGGTGA 1440  
Qy 1441 GATTTCATGACTGGGCTG 1457  
Db 1441 GATTTCATGACTGGGCTG 1457

RESULT 5  
ADM32723  
ID ADM32723 brandard; DNA; 1491 BP.

XX ADM32723;  
XX 17-JUN-2004 (first entry)  
XX AOB Bp16clone57 16S rDNA.  
XX 16S rDNA; ammonia oxidizing bacteria; AOB; ammonia; nitrite;  
XX aqueous environment; freshwater; seawater; aquarium; SB.  
XX Nitrosomonas aestuarii.  
XX WO2004026772-A2.  
XX 01-APR-2004.  
XX 10-SEP-2003; 2003WO-US028210.  
XX 19-SEP-2002; 2002US-0386217P.  
XX 19-SEP-2002; 2002US-0386218P.  
XX 19-SEP-2002; 2002US-0386219P.  
XX (AQUA-) AQUARIA INC.  
XX Hovaneer TA;  
XX WPI; 2004-304936/28.  
XX New composition comprising an isolated bacterial strain that oxidizes  
XX ammonia to nitrite, useful for alleviating or preventing the accumulation  
XX of ammonia in aqueous environment.  
XX Claim 1; Page 15; 98pp; English.  
XX This sequence represents a 16S rDNA sequence derived from an ammonia  
XX oxidizing bacteria (AOB). This sequence may be used in a composition  
XX which comprises an isolated bacterial strain that oxidizes ammonia to  
XX nitrite. The composition may be used for alleviating or preventing the  
XX accumulation of ammonia in a medium. The ammonia is reduced by at least

CC 30% when compared with a level of ammonia that would exist in the absence  
CC of the bacterial strain. The composition is useful for alleviating or  
CC preventing the accumulation of ammonia in aqueous environment, e.g. a  
CC freshwater or seawater aquarium.

SQ Sequence 1491 BP; 392 A; 323 C; 457 G; 319 T; 0 U; 0 Other;

Query Match 95.5%; Score 1391.6; DB 12; Length 1491;  
Best local similarity 97.6%; Pred. No. 0;  
Matches 1423; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Qy 1 ATTGAACGCTGCGGCATGCTTTTACATGCAATGCAAGTCGAACGCGACGATGCTTGAT 60  
Db 18 ATTGAACGCTGCGGCATGCTTTTACATGCAATGCAAGTCGAACGCGACGATGCTTGAT 77  
Qy 61 CTGTGGCGAGTGGCGGACGCGGTAGTATGCTATCGGAACCTATCCAGAAAGGGGGTAA 120  
Db 78 CTGTGGCGAGTGGCGGACGCGGTAGTATGCTATCGGAACCTATCCAGAAAGGGGGTAA 137  
Qy 121 ACGCATGAAAGATGTGCTTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAA 180  
Db 138 ACGCATGAAAGATGTGCTTAATACCGCATATATCTTAAGAGAGAAAGCAGGGGATCGAAA 197  
Qy 181 GACCTTGGCTTTTGAACGCGCGATGCTGATTAAGCTAGTATGCTAGTGGGTAAAGGCTTAC 240  
Db 198 GACCTTGGCTTTTGAACGCGCGATGCTGATTAAGCTAGTATGCTAGTGGGTAAAGGCTTAC 257  
Qy 241 CAAGCGACGATCAAGTATGCTGAGAGAGACCAACCGCACTGGGATGAGACAG 300  
Db 258 CAAGCGACGATCAAGTATGCTGAGAGAGACCAACCGCACTGGGATGAGACAG 317  
Qy 301 GCCCAGACTCTTACGCGGAGCAGCAGTGGGAAATTTTGAACAAATGGCGCAAGCCTGATC 360  
Db 318 GCCCAGACTCTTACGCGGAGCAGCAGTGGGAAATTTTGAACAAATGGCGCAAGCCTGATC 377  
Qy 361 CAGCAATGCGCGGTGAGTGAAGAAAGCCCTTGGGTTGTAAGCTCTTTCAGTGAAGAGA 420  
Db 378 CAGCAATGCGCGGTGAGTGAAGAAAGCCCTTGGGTTGTAAGCTCTTTCAGTGAAGAGA 437  
Qy 421 AAAGGTTACGTTAAATATATGCTGACTCATGACGCTATCGACAGAAAGACCCGGCTTAC 480  
Db 438 AAAGGTTGTTGCTTAATATATCAAACTTATGACGCTATCGACAGAAAGACCCGGCTTAC 497  
Qy 481 TACGTGCCAGAGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGGAATTAATGCTGGCGT 540  
Db 498 TACGTGCCAGAGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGGAATTAATGCTGGCGT 557  
Qy 541 AAAGGTTGCGCAGCGCGCTTTGTAAGTCAAGTGAATCCCGGGCTTAACTTGGGAAT 600  
Db 558 AAAGGTTGCGCAGCGCGCTTTGTAAGTCAAGTGAATCCCGGGCTTAACTTGGGAAT 617  
Qy 601 TCGCTTTGAACCTACAAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTTTACAG 660  
Db 618 TCGCTTTGAACCTACAAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTTTACAG 677  
Qy 661 TGAATGCGTGAAGTATGAAAGAACATCGATGCGGAAGGAGCGCTCTGGGTTAAACT 720  
Db 678 TGAATGCGTGAAGTATGAAAGAACATCGATGCGGAAGGAGCGCTCTGGGTTAAACT 737  
Qy 721 GACGCTCATGACAGAAACGTTGGGAGAGCAACAGATTTAGTACCTGTTAGTCCAGGCC 780  
Db 738 GACGCTCATGACAGAAACGTTGGGAGAGCAACAGATTTAGTACCTGTTAGTCCAGGCC 797  
Qy 781 CTTAAACGATGCAACTAGTGTGGGCTTTATGAGCTTGGTAAAGCAAGCTAAAGCGGTGA 840  
Db 798 CTTAAACGATGCAACTAGTGTGGGCTTTATGAGCTTGGTAAAGCAAGCTAAAGCGGTGA 857  
Qy 841 AGTTGACCGCTGGGAGTACGTCGCAAGATTAAATCTCAAGAAATTTGAACGGGAGCCC 900  
Db 858 AGTTGACCGCTGGGAGTACGTCGCAAGATTAAATCTCAAGAAATTTGAACGGGAGCCC 917  
Qy 901 GCACAAGCGGTGATTAATGTTGATTAATTCGATCAACGCGCAAAATCTTACCTTACCTT 960



Db	759	CTAAACGATGTCACTAGTTGTTGGGCTTCTACGGCTTGTAACGTACTAAGCGTCA	858
Qy	841	AGTTGACCGCTCTGGGAGTAGCGTGCAGATTTAAACTCAAGAAATTGACGGGAGCC	900
Db	859	AGTTGACCGCTCTGGGAGTAGCGTGCAGATTTAAACTCAAGAAATTGACGGGAGCC	918
Qy	901	GCACAAAGCGGTGGATTATGTGTGAATTAAATTGCATGCAAGCGGAAAACTTACCTACCTT	960
Db	919	GCACAAAGCGGTGGATTATGTGTGAATTAAATTGCATGCAAGCGGAAAACTTACCTACCTT	978
Qy	961	GACATGTAGCGAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTAACACAGGTGCT	1011
Db	979	GACATGTAGCGAATTTTCTAGAGATAGATTAGTGCTTCGGGAAACGCTAACACAGGTGCT	1033
Qy	1020	GCATGGCTGTGCTCAGCTCGTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGAAC	1077
Db	1039	GCATGGCTGTGCTCAGCTCGTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGAAC	1099
Qy	1080	CCTTGTCAATTAATTCGCATCATTTGGTGGGCACTTAATGAGATCGCGGTGACAAACC	1133
Db	1099	CCTTGTCAATTAATTCGCATCATTTAGTTGGGCACTTAATGAGATCGCGGTGACAAACC	1155
Qy	1140	GGAGGAAGGTGGGAGTAGAGTCAAGTCTTCATGCGCCCTTATGGGTAGGCGCTTCAACGTA	1199
Db	1159	GGAGGAAGGTGGGAGTAGAGTCAAGTCTTCATGCGCCCTTATGGGTAGGCGCTTCAACGTA	1211
Qy	1200	ATACAAATGGCGGTTACAGAGGGTTGCGCAACCGCGAGGGGAGACTAACTCAAGAAAGCGC	1255
Db	1219	ATACAAATGGCGGTTACAGAGGGTTGCGCAACCGCGAGGGGAGACTAACTCAAGAAAGCGC	1277
Qy	1260	GTCTGATGTCGGAGTCGGAATCTGCAACTCGACTCCGTGAAAGTCGGAATGCTTAGTATCG	1311
Db	1279	GTCTGATGTCGGAGTCGGAATCTGCAACTCGACTCCGTGAAAGTCGGAATGCTTAGTATCG	1333
Qy	1320	CGGATCAGCATGTCCCGGTGAATACGTTCCCGGCTTGTATACACACCGCGCGTCAACACA	1377
Db	1339	CGGATCAGCATGTCCCGGTGAATACGTTCCCGGCTTGTATACACACCGCGCGTCAACACA	1399
Qy	1380	TGGAGTGGGTTTCAACAGAAACAGGTAGTCTTAACCGTAAGGAGGGCGCTTGCAACGATG	1433
Db	1399	TGGAGTGGGTTTCAACAGAAACAGGTAGTCTTAACCGTGGAGGAGCGCTTGCAACGATG	1455
Qy	1440	AGATTCAATGACTGGGG	1455
Db	1459	TGCTCAATGACTTTGGG	1474

RESULT 7	
ADM32721	
ID	ADM32721 standard; DNA; 1467 BP.
XX	
XX	ADM32721;
AC	
XX	
DT	17-JUN-2004 (first entry)
XX	
DE	AOB PaClone42 16S rDNA.
XX	
KW	16S rDNA; ammonia oxidising bacteria; AOB; ammonia; nitrite;
XX	aqueous environment; freshwater; seawater; aquarium; ss.
XX	
OS	Nitrosomonas aestuarii.
XX	
PN	WO2004026772-A2.
XX	
PD	
XX	01-APR-2004.
XX	
PF	10-SEP-2003; 2003WO-US028210.
XX	
PR	19-SEP-2002; 2002US-0386217P.
PR	19-SEP-2002; 2002US-0386218P.
PR	19-SEP-2002; 2002US-0386219P.
XX	
PA	(AQUA-) AQUARIA INC.

XX	Hovaneec TA;
E1	
XX	WPI; 2004-304936/28.
DR	
XX	New composition comprising an isolated bacterial strain that oxidizes
PT	ammonia to nitrite, useful for alleviating or preventing the accumulation
PT	of ammonia in aqueous environment.
XX	
PS	Claim 1; Page 13; 98pp; English.
XX	
CC	This sequence represents a 16S rDNA sequence derived from an ammonia
CC	oxidizing bacteria (AOB). This sequence may be used in a composition
CC	which comprises an isolated bacterial strain that oxidizes ammonia to
CC	nitrite. The composition may be used for alleviating or preventing the
CC	accumulation of ammonia in a medium. The ammonia is reduced by at least
CC	30% when compared with a level of ammonia that would exist in the absence
CC	of the bacterial strain. The composition is useful for alleviating or
CC	preventing the accumulation of ammonia in aqueous environment, e.g. a
CC	freshwater or seawater aquatium.
CC	
XX	Sequence 1467 BP; 384 A; 320 C; 451 G; 312 T; 0 U; 0 Other;
SQ	
	Query Match            93.3%; Score 1358.8; DB 12; Length 1467;
	Best Local Similarity   97.2%; Pred. No. 0;
	Matches 1414; Conservative   0; Mismatches   37; Indels   3; Gaps   3
OY	1 ATGGAACGGCTGGCGGCATCTTTACATGCAAGTCGAAACGGCAGCATGCTTGCAT 60
Dd	16 ATTGAACGCTGGCGGCATCTTTACATGCAAGTCGAAACGGCAGCATGCTTGCAC 75
OY	61 CTGGTGCGAGATGGCGGAGCGGGTGAATATGCATCGAACGTATCCGAAGAAGGGGGTA 120
Dd	76 CTGGTGCGAGATGGCGGAGCGGGTGAATATGCATCGAACGTATCCGAAGTGGGGATA 135
OY	121 ACCGATCGAAGAATGTGCTAATACCGCATATCTTAAAGAGGAAACAGCGGATCGMA 180
Dd	136 ACCGATCGAAGAATGTGCTAATACCGCATATCTTAAAGAGGAAACAGGGATCGMA 195
OY	181 GACTTGGGCTTTTGGAGCGGCCGATGTCTAATTAGCATGTGGTGGGGTAAAGGCTTAC 240
Dd	196 GACTTGTGCTTTTGGAGCGGCCGATGTCTAATTAGCATGTGGTGGGGTAAAGGCTTAC 255
OY	241 CAAGCGCAGATCAGTAGTTGTGTCGAGAGGACGACACGACACTGGGACTGAGCACG 300
Dd	256 CAAGCGAAGATCAGTAGTTGTGTCGAGAGGACGACACGACACTGGGACTGAGCACG 315
OY	301 GCCCAGACTCTTACCGGAGGCGACAGTGGGGAAATTTTGAACAATGGCGCAAGCTGATC 360
Dd	316 GCCCAAGACTCTTACCGGAGGCGACAGTGGGGAAATTTTGAACAATGGCGCAAGCTGATC 375
OY	361 CAGCAATGCGGGGTAGAGGAAGAAGCCCTGGGGTGTAAAGCTCTTTCAGTGGAGAGA 420
Dd	376 CAGCAATGCGGGGTAGAGGAAGAAG-CTTGGGGTGTAAAGCTCTTTCAGTGGAGAGA 434
OY	421 AAAGGTTACGATTAATTAATCGTAGCTCATGACGGTATCGACAGAGAAGCACCGGCTAAC 480
Dd	435 AAAGGTTGACATTAATTAATCAACTTATGATGGTACCAGACAGAAGAAGCACCGGCTAAC 494
OY	481 TAGTGCCAGCAGCCCGGTAAATACGTAGGGGTGCAAGCGTTAATCGGAATTACTGGGCGT 540
Dd	495 TAGTGCCAGCAGCCCGGTAAATACGTAGGGGTGCAAGCGTTAATCGGAATTACTGGGCGT 554
OY	541 AAAAGGTGGCGAGGGGGCTTTGTAAAGTACGATGTGGAATCCCGGGGCTTAACTGGGAAT 600
Dd	555 AAAAGGTGGCGAGGGGGCTTTGTAAAGTACGATGTGGAATCCCGGGGCTTAACTGGGAAT 614
OY	601 TGCGTTTGAATCTACAAAGCTAGAGTGTGGAGAGGAGGTGGAATTCACATGTGTAGCAG 660
Dd	615 TGCGTTTGAATCTACAAAGCTAGAGTGTGGAGAGGAGGGGTGGAATTCATGTGTAGCAG 674
OY	661 TGAATATGCTAGAGTATGGAAGAACATGATGGCGGAAGGCGCTTCTGGGTTAACT 720
Dd	

Db 675 TGAATGCGTAGATATGAGAAACATGCGAAGCGACCCCTGGGTTAACT 734  
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAAGTATCCTGTAGTCCAGCC 780  
Db 735 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAAGTATCCTGTAGTCCAGCC 794  
QY 781 CTAACGATGTCAATAGTTGTTGGGCTTTATTAAGCTTGTAACGAGCTAACCGGTGA 840  
Db 795 CTAAACGATGTCAATAGTTGTTGGGCTTTATTAAGCTTGTAACGAGCTAACCGGTGA 854  
QY 841 AGTTGACCGCTGGGAGATACCGTCCCAAGATTAATACTCAAGAAATGACGGGAGCC 900  
Db 855 AGTTGACCGCTGGGAGATACCGTCCCAAGATTAATACTCAAGAAATGACGGGAGCC 914  
QY 901 GCAACGCGGTGATTAATGATGATTAATGATGCAACGCGAATACTTACCTACCTT 960  
Db 915 GCAACGCGGTGATTAATGATGATTAATGATGCAACGCGAATACTTACCTACCTT 974  
QY 961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC-TTCGGAAACGCTAACACAGGTCT 1019  
Db 975 GACATGTAGCGAATTTTCTAGAGATAGATTAGTGC-TTCGGAAACGCTAACACAGGTCT 1034  
QY 1020 GCATGCTGTCTGACGTCTGTCTGACGATGTTGGTTAAGTCCGCAACGAGCGCAAC 1079  
Db 1035 GCATGCTGTCTGACGTCTGTCTGACGATGTTGGTTAAGTCCGCAACGAGCGCAAC 1094  
QY 1080 CCTGTCAATTAATTTGCGCATGTTGGTGGGCACTTAATGAGACTGCGGTGACAAAC 1139  
Db 1095 CCTGTCAATTAATTTGCGCATGTTGGTGGGCACTTAATGAGACTGCGGTGACAAAC 1154  
QY 1140 GAGAGAAAGTGGGAGATGACGTCATGAGCCCTTAATGGGTAGGGCTTCAACAGTA 1199  
Db 1155 GAGAGAAAGTGGGAGATGACGTCATGAGCCCTTAATGGGTAGGGCTTCAACAGTA 1214  
QY 1200 ATTCATATGCGCGTACAGAGGTTGGCCCAACCCGAGAGGGGAGCTATCTCAAGAAACGC 1259  
Db 1215 ATTCATATGCGCGTACAGAGGTTGGCCCAACCCGAGAGGGGAGCTATCTCAAGAAACGC 1274  
QY 1260 GTGATGTCGCGATCGGAGTCTGCACTGCACTCCGTAAGTCCGTAATTCG 1319  
Db 1275 GTGATGTCGCGATCGGAGTCTGCACTGCACTCCGTAAGTCCGTAATTCG 1334  
QY 1320 CGGATCGCATGTCGCGGTGAATACGTTCCCGGCTTTGTAACACACCGCCGTCAACCA 1379  
Db 1335 CGGATCGCATGTCGCGGTGAATACGTTCCCGGCTTTGTAACACACCGCCGTCAACCA 1394  
QY 1380 TGGAGTGGGTTTCAACAGACGAGTATCTAACCTTAAGAGAGGGCGCTTGCACCGGTG 1439  
Db 1395 TGGAGTGGGTTTCAACAGACGAGTATCTAACCTTAAGAGAGGGCGCTTGCACCGGTG 1453  
QY 1440 AGATTCAATGACTGG 1453  
Db 1454 AGATTCAATGACTGG 1467

RESULT 8  
ABA02418  
ID ABA02418 standard; DNA, 1458 BP.  
XX  
AC ABA02418;  
XX  
XX 29-AUG-2003 (revised)  
DT 04-MAR-2002 (first entry)  
XX  
DE Type B ammonia-oxidizing bacterium 16S rRNA gene sequence, R3clones5.  
XX  
XX Type B; ammonia-oxidizing bacterium; AOB; nitrite; 16S rRNA gene;  
KM ribosomal RNA; aquarium; aquaculture; waste water treatment;  
XX bioremediation; ds.  
XX Nitrosomonadales.  
OS  
XX  
PN M0200190312-A1.

XX  
PD 29-NOV-2001.  
XX  
PF 17-MAY-2001; 2001WO-US016265.  
XX  
PR 19-MAY-2000; 2000US-00573684.  
XX  
PA (AQUA-) AQUARIA INC.  
XX  
PI Hovanec TA, Burrell PC;  
XX  
DR MPI; 2002-075367/10.  
XX  
PT New bacteria capable of oxidizing ammonia to nitrite, for preventing or  
PT alleviating the accumulation of ammonia in fresh water aquaria, seawater  
PT aquaria and waste water.  
XX  
PS Claim 2; Page 7; 62pp; English.  
XX  
CC The invention relates to 4 novel types of ammonia-oxidizing bacteria  
CC (AOB) found in freshwater aquaria. The bacteria are able to oxidise  
CC ammonia to nitrite and are members of the ammonia-oxidizing bacteria  
CC family of the beta subdivision of Proteobacteria. The 4 types of bacteria  
CC can be distinguished on the basis of their 16S rRNA (ribosomal RNA) gene  
CC sequences (ABA02416-ABA02419), and are classified as AOB type A (e.g.,  
CC R3clones140), type A1 (e.g., R3clones187), type B (e.g., R3clones5) and type  
CC C (e.g., R3clones47). The invention also encompasses isolated 16S rRNA  
CC gene sequences of the ammonia-oxidizing bacteria of the invention,  
CC oligonucleotide probes and primers for the detection of these bacteria,  
CC and compositions comprising the bacteria. The bacteria of the invention  
CC are useful in biological filters for reducing ammonia accumulation in  
CC both freshwater and seawater aquaria. They may also be used in waste  
CC water treatment and in bioremediation processes to reduce the level of  
CC pollution caused by ammonia. The present sequence represents R3clones5, a  
CC 16S rRNA gene sequence from the type B ammonia-oxidizing bacterium of the  
CC invention. (Updated on 29-AUG-2003 to standardise OS field)  
XX  
SQ Sequence 1458 BP; 368 A; 342 C; 466 G; 282 T; 0 U; 0 Other;  
Query Match 88.4%; Score 1287.6; DB 6; Length 1458;  
Best local similarity 93.6%; Pred. No. 0;  
Matches 1366; Conservative 0; Mismatches 89; Indels 5; Gaps 2;  
QY 1 ATTTGAACGCTGCGCGCATGCTTTTACATGCAAGTGAACCGGACGACCGGATCTTGCAAT 60  
Db 1 ATTTGAACGCTGCGCGCATGCTTTTACATGCAAGTGAACCGGACGACCGGATCTTGCAAT 58  
QY 61 CTGCTGGCGAGTGGCGGAGCGGTGAATATGATCGGAAAGTATCCAGAAAGAGGGGGGTA 120  
Db 59 CTGCTGGCGAGTGGCGGAGCGGTGAATATGATCGGAAAGTATCCAGAAAGAGGGGGGTA 118  
QY 121 ACCGATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGAGGGGATCGAAA 180  
Db 119 ACCGATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGAGGGGATCGAAA 178  
QY 181 GACCTTGGCGCTTTTGAAGCGCGCGATGTCTGATTAGCTAATGTTGGTGAAGGCTTAC 240  
Db 179 GACCTTGGCGCTTTTGAAGCGCGCGATGTCTGATTAGCTAATGTTGGTGAAGGCTTAC 238  
QY 241 CAAGGCGACATCAATGATTTGCTGAGAGACGACGACGACGACGACGACGACGACGACG 300  
Db 239 CAAGGCGACATCAATGATTTGCTGAGAGACGACGACGACGACGACGACGACGACGACG 298  
QY 301 GCCCAGACTCTCTACGCGGAGCAGACAGTGGGAAATTTTGAACAATGGCGGACGCTGATC 360  
Db 299 GCCCAGACTCTCTACGCGGAGCAGACAGTGGGAAATTTTGAACAATGGCGGACGCTGATC 358  
QY 361 CAGCAATGCGCGTGAAGTGAAGAGGCTTGGGTTGAAGCTTTTCAAGTGAAGAAGA 420  
Db 359 CAGCAATGCGCGTGAAGTGAAGAGGCTTGGGTTGAAGCTTTTCAAGTGAAGAAGA 418  
QY 421 AAAGTTAAGGTTAATTAATGATGATCATGACGCTATCGACAGAAAGACCGGCTTAC 480

Db 419 AAGGTCAGGCGTAATACCCGTGACTGACGCTACCGGAAGAAAGACACCGGCTAAC 478  
Qy 481 TACGTGCCAGCAGCCGCGGTAAATACGTAGGGTGCAAGCCTTAATCCGAATTACTGGCGCT 540  
Db 479 TACGTGCCAGCAGCCGCGGTAAATACGTAGGGTGCAAGCCTTAATCCGAATTACTGGCGCT 538  
Qy 541 AAAAGGTGGCGGGCGGCTTTGTAAGTCAGATGTGAAATCCCGGGCTTAACCTGGGAAT 600  
Db 539 AAAGCCTGCGCAGCGCGGTTTGTGAGTCAGATGTGAAACCCCGGGCTTAACCTGGGAAC 598  
Qy 601 TCGCTTTGAAACTACAGAGCTAGAGTGTGGCAGAGGAGGAGTGGAAATCCATGTGTAGCAG 660  
Db 599 TCGCTTTGAAACTACAGAGCTAGAGTGTGGCAGAGGAGGAGTGGAAATCCATGTGTAGCAG 658  
Qy 661 TGAATGCGTAGAGATATGAAAGAAATGATGCGAGGCGAGCCTCTGGGTTAACTACT 720  
Db 659 TGAATGCGTAGAGATATGAGAGAACACCGATGCGAGGCGAGCCTCTGGGTTAACTACT 718  
Qy 721 GACGCTCAGCAGCAAGAGGCGGGGAGCAACAGATTAATACCTGGTAGTCCACGCGC 780  
Db 719 GACGCTCAGCAGCAAGAGGCGGGGAGCAACAGATTAATACCTGGTAGTCCACGCGC 778  
Qy 781 CTAAACGATGTCAGTACTAGTTGTTGGGCTTATTAAGGCTTGTAACGAACTAACGCGTGA 840  
Db 779 CTAAACGATGTCAGTACTAGTTGTTGGGCTTATTAAGGCTTGTAACGCGTGA 838  
Qy 841 AGTTGACCGCTGGGGAGTACGCTGCGCAAGATTAATACTCAAGAAATTGACGGGGACCC 900  
Db 839 AGTTGACCGCTGGGGAGTACGCTGCGCAAGATTAATACTCAAGAAATTGACGGGGACCC 898  
Qy 901 GGAACAAGGCGGTGATTAATGATTAATTCGATGCAACGCGAAACCTTACTTACCTT 960  
Db 899 GGAACAAGGCGGTGATTAATGATTAATTCGATGCAACGCGAAACCTTACTTACCTT 958  
Qy 961 GACATGTACGCAATTTTCTAGAGATAGATTAGTGC--TTCCGGAACGCTTAACACAGGTG 1017  
Db 959 GACATGTACGCAAGCCCGCGGAGAGGTGGGTGTGCCGAAGGAGCGGTTAACACAGGTG 1018  
Qy 1018 CTGCATGCGCTGCTGCACTCGTGTGTAAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1077  
Db 1019 CTGCATGCGCTGCTGCACTCGTGTGTAAGATGTTGGGTTAAAGTCCCGCAACGAGCGCA 1078  
Qy 1078 ACCCTTGTCATTAATTGGCATGATTTGGTGGGCACTTTAAAGAGATGCGCGGTGACAA 1137  
Db 1079 ACCCTTGTCATTAATTGGCATGATTTGGGCACTTTAAAGAGATGCGCGGTGACAA 1138  
Qy 1138 CCGAGGAAGGTGGGGATGACGTCAAGTCTCATGCGCCTTAATGGGTAGGGCTTACACAG 1197  
Db 1139 CCGAGGAAGGTGGGGATGACGTCAAGTCTCATGCGCCTTAATGGGTAGGGCTTACACAG 1198  
Qy 1198 TAAATCAATGCGCGCTACAGAGGTTGCCAACCCCGAGGGGAGCTTAATCTCAGAAAGC 1257  
Db 1199 TAAATCAATGCGCGCTACAGAGGTTGCCAACCCCGAGGGGAGCTTAATCTCAGAAAGC 1258  
Qy 1258 GGTGTGATGTCGGAGTCGGAATCTGCAACTCGGAACTCGGAAATGCTGTATAT 1317  
Db 1259 GGTGTGATGTCGGAGTCGGAATCTGCAACTCGGAACTCGGAAATGCTGTATAT 1318  
Qy 1318 CCGCATCAGCATGTGCGCGGTAAATACGTTCCCGGGTCTTGTACACCCCGCTGACAC 1377  
Db 1319 CCGCATCAGCATGTGCGCGGTAAATACGTTCCCGGGTCTTGTACACCCCGCTGACAC 1378  
Qy 1378 CATGGAGTGGGTTTCAACGAAGAGGATGTCTAACCGTTAAAGAGGGGCTTTGCCACG 1437  
Db 1379 CATGGAGTGGGTTTCAACGAAGAGGATGTCTAACCGTTAAAGAGGGGCTTTGCCACG 1438  
Qy 1438 TGAAGTTCACTGACTGGGGTG 1457  
Db 1439 TGAAGTTCACTGACTGGGGTG 1458

RESULT 9  
ADM32706

ID ADM32706 standard; DNA; 1458 BP.  
XX  
AC ADM32706;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE AOB Type B R3clones 16S rDNA.  
XX  
KW 16S rDNA; ammonia oxidizing bacteria; AOB; ammonia; nitrite;  
XX aqueous environment; freshwater; seawater; aquarium; ss.  
XX  
OS Nitrosomonas sp.  
XX  
PN MO2004026772-A2.  
XX  
PD 01-APR-2004.  
XX  
PF 10-SEP-2003; 2003MO-US028210.  
XX  
PR 19-SEP-2002; 2002US-0386217P.  
XX  
PR 19-SEP-2002; 2002US-0386218P.  
XX  
PR 19-SEP-2002; 2002US-0386219P.  
XX  
PA (AQUA-) AQUARIA INC.  
XX  
PI Hovanez TA.  
XX  
DR WPI; 2004-304936/28.  
XX  
PT New composition comprising an isolated bacterial strain that oxidizes  
PT ammonia to nitrite, useful for alleviating or preventing the accumulation  
PT of ammonia in aqueous environment.  
XX  
PS Disclosure; Page 11; 98pp; English.  
XX  
CC This sequence represents a 16S rDNA sequence derived from an ammonia  
CC oxidizing bacteria (AOB). This sequence may be used in a composition  
CC which comprises an isolated bacterial strain that oxidizes ammonia to  
CC nitrite. The composition may be used for alleviating or preventing the  
CC accumulation of ammonia in a medium. The ammonia is reduced by at least  
CC 30% when compared with a level of ammonia that would exist in the absence  
CC of the bacterial strain. The composition is useful for alleviating or  
CC preventing the accumulation of ammonia in aqueous environment, e.g. a  
CC freshwater or seawater aquarium.  
XX  
SQ Sequence 1458 BP; 368 A; 342 C; 466 G; 282 T; 0 U; 0 Other;  
Query Match 88.4%; Score 1287.6; DB 12; Length 1458;  
Best Local Similarity 93.6%; Pred. No. 0;  
Matches 1366; Conservative 0; Mismatches 89; Indels 5; Gaps 2;  
Qy 1 ATTGAACGCTGCGGCGATGCTTTACATGCAAGTGCAGACGCGACGATGCTTGCAAT 60  
Db 1 ATTGAACGCTGCGGCGATGCTTTACATGCAAGTGCAGACGCGACGCGATGCTTGCAAT 58  
Qy 61 CTGTGTGAGATGTGGCGGCGGGTGAATGATCGGAACGATTCAGAAAGAGGGGGGTA 120  
Db 59 CTGTGTGAGATGTGGCGGCGGGTGAATGATCGGAACGATTCAGAAAGAGGGGGGTA 118  
Qy 121 ACGCATCGAAAGATGTGTAATACCGCATATCTTAAGAGAAAGCAGGGGATCGAAA 180  
Db 119 ACGCATCGAAAGATGTGTAATACCGCATATCTTAAGAGAAAGCAGGGGATCGCAA 178  
Qy 181 GACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAAGTGAATTTGGTGGGTTAAAGCCTTAC 240  
Db 179 GACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAAGTGAATTTGGTGGGTTAAAGCCTTAC 238  
Qy 241 CAAGCGGACGATTCAGTACTGTTGCTGAGAGGACGACGACGCTGGGACTGAGACAG 300  
Db 239 CAAGCGGACGATTCAGTACTGTTGCTGAGAGGACGACGACGCTGGGACTGAGACAG 298  
Qy 301 GCCCAGACTCTTACGGGAGCAGAGTGGGGAATTTTGAACATGCGCGCAAGCCTGATC 360

Db 299 GCCCAGACTCTTACGCGGAGGCGAGTGGGGAATTTGGACATGCGGGGAAACCTGATC 358  
 Qy 361 CAGCAATGCGCGGTGAGTGAAGAAAGCCCTTGGGTTGTAAGCTCTTTCAGTCGAGAGA 420  
 Db 359 CAGCAATGCGCGGTGAGTGAAGAAAGCCCTTGGGTTGTAAGCTCTTTCAGTCGAGAGA 418  
 Qy 421 AAAGTTACGTTAAATTAATCGTACTCATGACGGTATCGACAGAAAGAACCGGCTAAC 480  
 Db 419 AACGGTCAACGGTAAATACCGGTACTGACGGTACCGGAAAGAACCGGCTAAC 478  
 Qy 481 TAGCTGCCAGCGCGCGGTATAGTGAAGGTGCGAAGCTTATGGAATTAATGCGGCGT 540  
 Db 479 TAGCTGCCAGCGCGCGGTATAGTGAAGGTGCGAAGCTTATGGAATTAATGCGGCGT 538  
 Qy 541 AAAGGTGCGAGCGCGCTTGTAGTCAAGTGTGAATCCCGGCGCTTAACTGGGAAT 600  
 Db 539 AAAGGTGCGAGCGCGCTTGTAGTCAAGTGTGAATCCCGGCGCTTAACTGGGAAT 598  
 Qy 601 TCGGTTTGAACCTACAAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660  
 Db 599 TCGGTTTGAACCTACAAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 658  
 Qy 661 TGAATGCTAGATATGGAAGACATGATGCGAAGCGCTCTCGGTTAACT 720  
 Db 659 TGAATGCTAGATATGGAAGACATGATGCGAAGCGCTCTCGGTTAACT 718  
 Qy 721 GACGCTCAGCAGAAAGGCGGAGCAAAACAGATTGATCCCTGGTATGCAAGCC 780  
 Db 719 GACGCTCAGCAGAAAGGCGGAGCAAAACAGATTGATCCCTGGTATGCAAGCC 778  
 Qy 781 CTAAACGATGTCACTAGTGTGTGGCTTATATGAGCTTGTATACGAAGCTAACCGGTGA 840  
 Db 779 CTAAACGATGTCACTAGTGTGTGGCTTATATGAGCTTGTATACGAAGCTAACCGGTGA 838  
 Qy 841 AGTTGACCGCTGGGAGGATCGGTGCGAAGATTAACTCAAGAAATTGAACGGGAGCC 900  
 Db 839 AGTTGACCGCTGGGAGGATCGGTGCGAAGATTAACTCAAGAAATTGAACGGGAGCC 898  
 Qy 901 GCAACAGCGGTGATTTATGCGATTATTCGATGCAACGCGGAAACCTTACCTACCTT 960  
 Db 899 GCAACAGCGGTGATTTATGCGATTATTCGATGCAACGCGGAAACCTTACCTACCTT 958  
 Qy 961 GACATGTAGCAATTTTCTAGATAGTATGATGCTGCTGCGGAAACGCTTACAGAGGT 1017  
 Db 959 GACATGTAGCAATTTTCTAGATAGTATGATGCTGCTGCGGAAACGCTTACAGAGGT 1018  
 Qy 1018 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
 Db 1019 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078  
 Qy 1078 ACCCTGTCTTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
 Db 1079 ACCCTGTCTTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1138  
 Qy 1138 CCGGAGGAGGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGAGGTAGGCTTCAACG 1197  
 Db 1139 CCGGAGGAGGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGAGGTAGGCTTCAACG 1198  
 Qy 1198 TAAATCAATGCGCGTACAGAGGCTTCCAAACCGCGGAGGAGGAGTAACTTCAAGAAC 1257  
 Db 1199 TAAATCAATGCGCGTACAGAGGCTTCCAAACCGCGGAGGAGGAGTAACTTCAAGAAC 1258  
 Qy 1258 GCGCTGTAGTCCGAGTCTGCAACTGCACTCGCTTAAAGTCCGAAATCGCTAGTAAT 1317  
 Db 1259 GCGCTGTAGTCCGAGTCTGCAACTGCACTCGCTTAAAGTCCGAAATCGCTAGTAAT 1318  
 Qy 1318 CCGGATCAGCATGCGCGGTATAGTCCCGGCTTGTATACACACCGCGCTTCAAC 1377  
 Db 1319 CCGGATCAGCATGCGCGGTATAGTCCCGGCTTGTATACACACCGCGCTTCAAC 1378  
 Qy 1378 CATGGAGTGGGTTTCAACAGAGCAGGTACTTAACTGTAAGAGAGGCGCTTCCACGG 1437  
 Db 1379 CATGGAGTGGGTTTCAACAGAGCAGGTACTTAACTGTAAGAGAGGCGCTTCCACGG 1438

Qy 1438 TGAGATTGATGATCTGGGTG 1457  
 Db 1439 TGAGATTGATGATCTGGGTG 1458  
 RESULT 10  
 ABA02419  
 ID ABA02419 standard; DNA; 1460 BP.  
 XX  
 AC ABA02419;  
 XX  
 DT 29-AUG-2003 (revised)  
 DT 04-MAR-2002 (first entry)  
 XX  
 DE Type C ammonia-oxidising bacterium 16S rRNA gene sequence, R3clone47.  
 XX  
 KW Type C; ammonia-oxidising bacterium; AOB; nitrite; 16S rRNA gene;  
 KW ribosomal RNA; aquarium; aquaculture; waste water treatment;  
 KW bioremediation; ds.  
 XX  
 OS Nitrosomonadales.  
 XX  
 PN WO200190312-A1.  
 PD 29-NOV-2001.  
 XX  
 PF 17-MAY-2001; 2001WO-US016265.  
 XX  
 PR 19-MAY-2000; 2000US-00573684.  
 XX  
 PA (AQUA-) AQUARIA INC.  
 XX  
 PI Hovanec TA, Burrell PC;  
 DR WPI; 2002-075367/10.  
 XX  
 PT New bacteria capable of oxidizing ammonia to nitrite, for preventing or  
 PT alleviating the accumulation of ammonia in fresh water aquaria, seawater  
 PT aquaria and waste water.  
 PS Claim 2; Page 8; 62pp; English.  
 XX  
 CC The invention relates to 4 novel types of ammonia-oxidising bacteria  
 CC (AOB) found in freshwater aquaria. The bacteria are able to oxidise  
 CC ammonia to nitrite and are members of the ammonia-oxidising bacteria  
 CC family of the beta subdivision of Proteobacteria. The 4 types of bacteria  
 CC can be distinguished on the basis of their 16S rRNA (ribosomal RNA) gene  
 CC sequences (ABA02416-ABA02419), and are classified as AOB type A (e.g.,  
 CC R3clone140), type A1 (e.g., R3clone187), type B (e.g., R3clone5) and type  
 CC C (e.g., R3clone47). The invention also encompasses isolated 16S rRNA  
 CC gene sequences of the ammonia-oxidising bacteria of the invention,  
 CC oligonucleotide probes and primers for the detection of these bacteria,  
 CC and compositions comprising the bacteria. The bacteria of the invention  
 CC are useful in biological filters for reducing ammonia accumulation in  
 CC both freshwater and seawater aquaria. They may also be used in waste  
 CC water treatment and in bioremediation processes to reduce the level of  
 CC pollution caused by ammonia. The present sequence represents R3clone47, a  
 CC 16S rRNA gene sequence from the type C ammonia-oxidising bacterium of the  
 CC invention. (Updated on 29-AUG-2003 to standardise OS field)  
 CC  
 XX  
 XX Sequence 1460 BP; 388 A; 316 C; 454 G; 302 T; 0 U; 0 Other;  
 Query Match 84.4%; Score 1230.2; DB 6; Length 1460;  
 Best Local Similarity 92.3%; Pred. No. 0;  
 Matches 1351; Conservative 0; Mismatches 103; Indels 9; Gaps 5;  
 Qy 1 ATTGAACGCTGCGCGGATGCTTATACATGCAAGTGAACGCGCAGCAGGATCTTGCAT 60  
 Db 1 ATTGAACGCTGCGCGGATGCTTATACATGCAAGTGAACGCGCAGG--CGGCGCTTCGCG 58  
 Qy 61 CTGCTGCGCAGTGGCGGAGCGGTGATTAATGATTCGGAACGTATCCGAAGAGGCGGCTA 120  
 Db 61 CTGCTGCGCAGTGGCGGAGCGGTGATTAATGATTCGGAACGTATCCGAAGAGGCGGCTA 120



Db 59 CTGCCGCGAGTGGCGAAGCGGTGAGTATATACATCGAAGCGTCTTAAAGTGGGAAATA 118  
Qy 121 ACGCATCGAAAGATGTGCTAATAATCCGCATATATCTTAAGAGAGAAAAGCAGGGGATCGAAA 180  
Db 119 ACGCATCGAAAGATGTGCTAATAATCCGCATAT-CTCTGAGAGAGAAAAGCAGGGGATCGCAA 177  
Qy 181 GACCTTGGCGCTTTTGGAGCGCGGAGATGTGATTTAGCTAGTTGGTGGGCTTAAAGGCTTAC 240  
Db 178 GACCTTGGCGCTTAAAGAGCGCGGAGATGTGATTTAGCTAGTTGGTGGGCTTAAAGGCTTAC 237  
Qy 241 CAAGCGCAGCATCAGTATGTGTCTGAGAGAGACACAGCCACACTGGAGCTGAGACAG 300  
Db 238 CAAGCGCAGCATCAGTATGTGTCTGAGAGAGACACACACTGGAGCTGAGACAG 297  
Qy 301 GCCCAAGCTCTTACCGGAGGAGCAGAGTGGGAAATTTTGGACATATGGGCGCAAGCTGATC 360  
Db 298 GCCCAAGCTCTTACCGGAGGAGCAGAGTGGGAAATTTTGGACATATGGGCGCAAGCTGATC 357  
Qy 361 CAGCAATGCCGCTGATGTAAGAGGCGCTTGGGTTGTAAAGCTCTTTCAGTCAAGAAAGA 420  
Db 358 CAGCAATGCCGCTGATGTAAGAGGCGCTTGGGTTGTAAAGCTCTTTCAGTCAAGAAAGA 417  
Qy 421 AAAGGTTACGTTAAATATCTGACTCATGACGCTATCGACAGAAAGAACACCGGCTAAC 480  
Db 418 AAAGGTTACGTTAAATATCTGACTCATGACGCTATCGACAGAAAGAACACCGGCTAAC 477  
Qy 481 TACGTGCCAGCAGCCGCGGTAAATACGTAGGGTTCAGAGCTTATCGGAATTTCTGGGCGT 540  
Db 478 TACGTGCCAGCAGCCGCGGTAAATACGTAGGGTTCAGAGCTTATCGGAATTTCTGGGCGT 537  
Qy 541 AAAGGTTGGCAGGGGCGCTTGTAAAGTCAAGTGTGAAATCCCCGGGCTTAACTGGGAAT 600  
Db 538 AAAGGTTGGCAGGGGCGCTTGTAAAGTCAAGTGTGAAATCCCCGGGCTTAACTGGGAAT 597  
Qy 601 TCGCTTTGAAACTACAGGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGCAG 660  
Db 598 TCGCTTTGAAACTACAGGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGCAG 657  
Qy 661 TGAATTCGCTAGAGATATGAAAGACATGATGGCGAAGGCGCTCTGGGTTAACTACT 720  
Db 658 TGAATTCGCTAGAGATATGAAAGACATGATGGCGAAGGCGCTCTGGGTTAACTACT 717  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGAGAAACAGGATTGATACCTGGTATGCCAGCGC 780  
Db 718 GACGCTCATGACGAAAGCGTGGGAGAGAAACAGGATTGATACCTGGTATGCCAGCGC 777  
Qy 781 CTAAACGATGTCACTAGTTGTTGGGCTTATTA--GGCTTGTATACGAAAGCTTAACGCGT 838  
Db 778 CTAAACGATGTCACTAGTTGTTGGGCTTATTA--GGCTTGTATACGTAAGCTTAACGCGT 837  
Qy 839 GAAAGTTGACCGCTGGGAGATACGCTCGCAAGATTAAACCTCAAGAGAAATTGAAGCGGAC 898  
Db 838 GAAAGTTGACCGCTGGGAGATACGCTCGCAAGATTAAACCTCAAGAGAAATTGAAGCGGAC 897  
Qy 899 CCGCAAGCGGCTGATTAATGAGATTATTCGATGCAACGCGAAACCTTAACCTACCC 958  
Db 898 CCGCAAGCGGCTGATTAATGAGATTATTCGATGCAACGCGAAACCTTAACCTACCC 957  
Qy 959 TTGACATGATAGCAATTTTCTAGAGATAGATTAGTGC--TTCCGGGAAACGCTAACACAGG 1015  
Db 958 TTGACATGATGATAGCAATTTTCTAGAGATAGATTAGTGC--TTCCGGGAAACGCTAACACAGG 1017  
Qy 1016 TGTCTCATGCTGTGTGCTGACGCTGTGTGCTGAGATGTTGGGTTAAAGTCCCGCAAGAGCG 1075  
Db 1018 TGTCTCATGCTGTGTGCTGACGCTGTGTGCTGAGATGTTGGGTTAAAGTCCCGCAAGAGCG 1077  
Qy 1076 CAACCTTGTCTAATTAATGCGCATCAAT--TGTTGGGCACTTTAAAGAGATCGCGGTGAC 1134  
Db 1078 CAACCTTGTCTAATTAATGCGCATCAATTTAAAGAGACATTTAAGAGATCGCGGTGAC 1137  
Qy 1135 AAACCGGAGAGATGCGGATGACGTCAAGTCTCATGCGCTTATGAGGTAGGCGCTTAC 1194  
Db 1138 AAACCGGAGAGATGCGGATGACGTCAAGTCTCATGCGCTTATGAGGTAGGCGCTTAC 1197

Qy 1195 ACGTAATCAATATGCGCGCTACAGAGGCTTGGCAACCCGCGAGGGGGAGCTAATCTCAGAA 1254  
Db 1198 ACGTAATCAATATGCGCGCTACAGAGGCTTGGCAACCCGCGAGGGGGAGCTAATCTCAGAA 1257  
Qy 1255 AGCGCGTGTATGTCGAGATCGGATCTGCAACTGCACTCGCTGAAGTGGAAATGCTAGT 1314  
Db 1258 AGCACTGTGTATGTCGAGATCGGATCTGCAACTGCACTCGCTGAAGTGGAAATGCTAGT 1317  
Qy 1315 AATCGCGATCAAGCATGTGCGGCTGATATGCTTCCCGGGCTTGTACACACCGCCGTCA 1374  
Db 1318 AATCGCGATCAAGCATGTGCGGCTGATATGCTTCCCGGGCTTGTACACACCGCCGTCA 1377  
Qy 1375 CACCATGGAGTGTGCTTACACAGAGAGGTTAGCTTAAACGTAAGAGAGGCGCTTGCA 1434  
Db 1378 CACCATGGAGTGTGCTTACACAGAGAGGTTAGCTTAAACGTAAGAGAGCGCTTGCA 1437  
Qy 1435 CGGTGAGATTCATGACTGGGGTG 1457  
Db 1438 CGGTGAGGCTCATGACTGGGGTG 1460

RESULT 11  
ADM32707  
ID ADM32707 standard; DNA; 1460 BP.  
XX  
AC ADM32707;  
XX  
XX 17-JUN-2004 (first entry)  
DT  
XX  
DE AOB Type C R5c1one47 16S rDNA.  
XX  
KM 16S rDNA; ammonia oxidising bacteria; AOB; ammonia; nitrite;  
KM aqueous environment; freshwater; seawater; aquarium; ss.  
OS Nitrosomonas sp.  
XX  
PN WO2004026772-A2.  
XX  
PD 01-APR-2004.  
XX  
PF 10-SEP-2003; 2003WO-US028210.  
XX  
PR 19-SEP-2002; 2002US-0386219P.  
PR 19-SEP-2002; 2002US-0386219P.  
PR 19-SEP-2002; 2002US-0386219P.  
PA (AQUA-) AQUARIA INC.  
XX  
PI Hovaneec TA;  
XX  
DR WPI; 2004-304936/28.  
XX  
PT New composition comprising an isolated bacterial strain that oxidizes  
PT ammonia to nitrite, useful for alleviating or preventing the accumulation  
PT of ammonia in aqueous environment.  
XX  
PS Disclosure; Page 12; 98pp; English.  
XX  
CC This sequence represents a 16S rDNA sequence derived from an ammonia  
CC oxidising bacteria (AOB). This sequence may be used in a composition  
CC which comprises an isolated bacterial strain that oxidizes ammonia to  
CC nitrite. The composition may be used for alleviating or preventing the  
CC accumulation of ammonia in a medium. The ammonia is reduced by at least  
CC 30% when compared with a level of ammonia that would exist in the absence  
CC of the bacterial strain. The composition is useful for alleviating or  
CC preventing the accumulation of ammonia in aqueous environment, e.g. a  
CC freshwater or seawater aquarium.  
XX  
SQ Sequence 1460 BP; 388 A; 316 C; 454 G; 302 T; 0 U; 0 Other;

Query Match 84.4%; Score 1230.2; DB 12; Length 1460;  
Best Local Similarity 92.3%; Pred. No. 0;



Matches 1351; Conservative 0; Mismatches 103; Indels 9; Gaps 5;

```

Qy 1 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAACGGGACGCGATGCTTGCAT 60
Db 1 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAACGGGACG--CGGGGCTTCCGCG 58
Qy 61 CTGTGGCGAGTGGCGGACGGGTAGTAAATGCATCGGAACGTAATCCAAAGAGGGGGTAA 120
Db 59 CTGTGGCGAGTGGCGGACGGGTAGTAAATGCATCGGAACGTAATCCAAAGAGGGGGTAA 118
Qy 121 AGGCATCGAAAGATGTCTAATATCCGCATATATCTTAAGAGGAAAGCGGGATCGA 180
Db 119 AGGCATCGAAAGATGTCTAATATCCGCATATATCTTAAGAGGAAAGCGGGATCGA 177
Qy 181 GACCTGGCGCTTTTGGAGCGGCGGATGTCTGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 240
Db 178 GACCTGGCGCTTTAAAGAGCGGCGGATGTCTGATTAGCTAGTTGGTGGGTTAAAGGCTTAC 237
Qy 241 CAAGCGACGATCATAGTATGTGTCTGAGAGACGACCAAGCACAATGGGAATGAGACAG 300
Db 238 CAAGCGACGATCATAGTATGTGTCTGAGAGACGACCAACCACTGGGACTGAGACAG 297
Qy 301 GCCCAGACTCTTACGGGAGGCGACAGTGGGGAATTTTGGACATGGGGCGGAACCTGATC 360
Db 298 GCCCAGACTCTTACGGGAGGCGACAGTGGGGAATTTTGGACATGGGGCGGAACCTGATC 357
Qy 361 CAGCAATGCCGGGTGAGTGAAGAAAGGCTTCGGGTTGTAAGGCTTTTCGAGAGAGA 420
Db 358 CAGCAATGCCGGGTGAGTGAAGAAAGGCTTCGGGTTGTAAGGCTTTTCGAGAGAGA 417
Qy 421 AAAGTTACGCTTAATATATCTGATCATGACGCTATCGACAGAAAGACGCGCTAC 480
Db 418 AAAGTTACGCTTAATATATCTGATCATGACGCTATCGACAGAAAGACGCGCTAC 477
Qy 481 TACGTGCCAGACGCGCGGTTAATCTGATGGTGCAGACCTTAATCGAATTTACTGGGCGT 540
Db 478 TACGTGCCAGACGCGCGGTTAATCTGATGGTGCAGACCTTAATCGAATTTACTGGGCGT 537
Qy 541 AAAGGGTGGCGAGGCGGCTTGTAAAGTCAATGTAAGAAATCCCGGGCTTAACCTGGGAAT 600
Db 538 AAAGGGTGGCGAGGCGGCTTGTAAAGTCAATGTAAGAAATCCCGGGCTTAACCTGGGAAT 597
Qy 601 TGGCTTGAATCAACAGGCTAGAGTGTGCAAGGGAGGTGAATTCATGTGTAGCAG 660
Db 598 TGGCTTGAATCAACAGGCTAGAGTGTGCAAGGGAGGTGAATTCATGTGTAGCAG 657
Qy 661 TGAATGCGTAGATATGGAAGAACATCGATGGCGAAGCGCTCTCGGCTTAACTACT 720
Db 658 TGAATGCGTAGATATGGAAGAACATCGATGGCGAAGCGCTCTCGGCTTAACTACT 717
Qy 721 GAGGCTCATGACGAAGACGCTGGGAGACAAACAGGATTAATACCTGGTATGCAAGCC 780
Db 718 GAGGCTCATGACGAAGACGCTGGGAGACAAACAGGATTAATACCTGGTATGCAAGCC 777
Qy 781 CTAAACGATGCACTAGTTGTTGGGCTTATTA--GGCTTGTGAAGAAAGCTAACCGT 838
Db 778 CTAAACGATGCACTAGTTGTTGGGCTTATTA--GGCTTGTGAAGAAAGCTAACCGT 837
Qy 839 GAACTTGAACCGCTGGGAGAGTACGCTGCAAGATTAATACTCAAGAAATTTGACGGGAC 898
Db 838 GAACTTGAACCGCTGGGAGAGTACGCTGCAAGATTAATACTCAAGAAATTTGACGGGAC 897
Qy 899 CCGGACGAAGCGGTGATTAATGTGATTAATTCGATGCAAGCGGAAAAACCTTACTCC 958
Db 898 CCGGACGAAGCGGTGATTAATGTGATTAATTCGATGCAAGCGGAAAAACCTTACTCC 957
Qy 959 TTGACATGACGCAATTTCTAGAGATAGATTAGTGC---TTCCGGAACGCTAACACAGG 1015
Db 958 TTGACATGCTTGGAAATCTAGTGAAGACATAAGAGTGGCCGGAAGGAGCCAAACACAGG 1017
Qy 1016 TGTGCAATGCTGTGCTGCAAGTGTGCTGAGATGTTGGGTTAAAGTCCCGCAACGAGCG 1075
Db 1018 TGTGCAATGCTGTGCTGCAAGTGTGCTGAGATGTTGGGTTAAAGTCCCGCAACGAGCG 1077

```

```

Qy 1076 CAACCTTGTCACTTAATTTGCCATCATTT--TGGTGGGCACTTTAATGACTGCCGATGAC 1134
Db 1078 CAACCTTGTCACTTAATTTGTCTATCTTAATTAATGACCTTTAGTAGACTGCCGATGAC 1137
Qy 1135 AAACCGGAGGAAGTGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTAC 1194
Db 1138 AAACCGGAGGAAGTGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTAC 1197
Qy 1195 AGGTAAATCAATGCGCGCTGACAGAGGTTTCCCAACCCCGAGGGGAGCTAATCTCAGAA 1254
Db 1198 AGGTAAATCAATGCGCGTGTACAGAGGTTTCCCAACCCCGAGGGGAGCTAATCTCAGAA 1257
Qy 1255 AGCGGTGTCTATCCGGATTCGGAATCTGCAACTGCACTCCGGAATGGGAATGCTAGT 1314
Db 1258 AGCGGTGTCTATCCGGATTCGGAATCTGCAACTGCACTCCGGAATGGGAATGCTAGT 1317
Qy 1315 AATCGCGATCGACATGTGCGGCTGAATTCGTTCCGGGCTTTGTACACACCGCCGTCA 1374
Db 1318 AATCGCGATCGACATGTGCGGCTGAATTCGTTCCGGGCTTTGTACACACCGCCGTCA 1377
Qy 1375 CACCATGGGAGTGGGTTTTCACCAAGACAGTATCTTAACGTAAGAGAGGGCGTTGCCA 1434
Db 1378 CACCATGGGAGTGGGTTTTCACCAAGACAGTATCTTAACGTAAGAGAGGGCGTTGCCA 1437
Qy 1435 CGGTGAGATTCAATGACTGGGGTG 1457
Db 1438 CGGTGAGATTCAATGACTGGGGTG 1460

```

RESULT 12  
AAC86030  
ID AAC86030 standard; cDNA; 1485 BP.

AAC86030;  
AC AAC86030;  
XX  
DT 11-SRP-2003 (revised)  
DT 29-AUG-2001 (first entry)  
XX  
DE 16S rDNA, Propel.  
XX  
KW 16S rDNA, polyphosphate accumulating organism; PAO; probe; primer;  
KW detection; phosphorus; waste water; sludge; ss.  
XX  
OS Propionivibrio pelophilus.  
XX  
PN WO200146459-A1.  
XX  
PD 28-JUN-2001.  
XX  
PF 28-DEC-2000; 2000WO-AU001611.  
XX  
PR 23-DEC-1999; 99AU-00004867.  
XX  
PA (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.  
XX  
PI Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ;  
XX  
DR WPI; 2001-408656/43.  
XX  
XX The sequences given in AAC86021-30 represent 16S rDNA sequences from  
CC polyphosphate accumulating organisms (PAOs). Sequences which are unique  
CC to these 16S rDNA sequences are used to create a probe or primer for  
CC detecting the relevant organisms. The primer/probe sequences are useful  
CC for detecting PAO cells in a sample. By treating cells in the sample to  
CC fix cellular contents, contacting fixed cells with the primer/probe which  
CC is labelled with a radiolabel, a reporter group or a hapten, under

conditions which allow the probe to hybridize with 16S rRNA within the fixed cell, removing unhybridized probe from the fixed cells, and detecting the labeled probe-RNA hybrid by fluorescence in situ hybridization. The primer/probe sequences are useful for identifying PAOs that are capable of biologically removing phosphorus from waste water. Rapid assessment of the presence of a number of PAOs in a waste water sample, can be done using the primer/ probe sequences. They allow quick and convenient assessment of whether a sludge or waste water sample includes PAOs and allows quantitation of PAO cells in samples. (Updated on 11-SEP-2003 to standardise OS field)

Sequence 1485 BP; 367 A; 341 C; 482 G; 295 T; 0 U; 0 Other;

Query Match 82.8%; Score 1206.4; DB 4; Length 1485;  
Best Local Similarity 90.0%; Pred. No. 0;  
Matches 1316; Conservative 0; Mismatches 141; Indels 5; Gaps 2;

```
QY 1 ATTGAACGCTGGCGGATGCTTTACATGCAAGTGAACGGGAGACCGGATGCTTGCAAT 60
DB 8 ATTGAACGCTGGCGGATGCTTTACATGCAAGTGAACGGGAGATGGGTGCTTGAC 67
QY 61 CTGTGGCGAGTGGCGGACGGGTGAGTAAATGCAATCGGAACGTATCCAGAAAGGGGGATA 120
DB 68 CTGATGGCGAGTGGCGGACGGGTGAGTAAATGCAATCGGAACGTATCCGGAAGTGGGGGATA 127
QY 121 ACGCATCGAAAGATGCTTAATACCGCATATCTTAAGAGAGAAAGCGGGGATCGAAA 180
DB 128 ACGTACCGAAAGTAAAGCTTAATACCGCATATCTTAAGAGAGAAAGGGGGATCGCAA 187
QY 181 GACCTTGCGCTTTTGGAGCGCGCATGTCGATTTAGCTAGTGGGTAAAGGCTTAC 240
DB 188 GACCTTGCGCTTTTGGAGCGCGCATGTCGATTTAGCTAGTGGGTAAAGGCTTAC 247
QY 241 CAAGCGACGATCAGTAGTTGGTCTGAGAGAGACGACACCACTCGGACCTGAGACAG 300
DB 248 CAAGCGACGATCAGTAGTTGGTCTGAGAGATGATCCGCCACACTGGGACTGAGACAG 307
QY 301 GCCCGACTCTTACCGGAGGCGACGAGTGGGGAAATTTTGGACAATGGGCCCAAGCTGATC 360
DB 308 GCCCGACTCTTACCGGAGGCGACGAGTGGGGAAATTTTGGACAATGGGCCCAAGCTGATC 367
QY 361 CAGCATGCGCGGCTGAGTGAAGAGGCTTCGGGTGTTAAAGCTCTTTGAGTGAAGAGA 420
DB 368 CAGCATGCGCGGCTGAGTGAAGAGGCTTCGGGTGTTAAAGCTCTTTGAGTGAAGAGA 427
QY 421 AAAGGTTACGTAATATATGTAATCATGACCGTATCGACAGAAAGACACCGGCTAAC 480
DB 428 AATGGACGCTTAACATATGCGTGTGATGACGGTACCGACATTAAGAACACCGGCTAAC 487
QY 481 TACGTGCCAGACCGCGGTAATACGTAGGGTGCAGCGCTTAATCGGAATTAATGAGCGT 540
DB 488 TACGTGCCAGACCGCGGTAATACGTAGGGTGCAGCGCTTAATCGGAATTAATGAGCGT 547
QY 541 AAAGGTTGCGAGCGGCTTTGTAAGTCAAGTGAATCCCGGGCTTAACCTGAGGAAT 600
DB 548 AAAGGTTGCGAGCGGCTTTGTAAGTCAAGGTTGAATCCCGGGCTTAACCTGAGGAAT 607
QY 601 TCGCTTTGAAACTACAAGGCTAGAGTGTGCGACAGAGGAGTGGAAATTCATGTGACAG 660
DB 608 GGCCTTTGAGATGCGACCGGCTAGAGTGTGACAGAGGAGGAGTGAATTCACGATGACAG 667
QY 661 TGAATATGCTAGATATGGAAGACATGATGGCGAAGCGCTCTCGGGTAACTACT 720
DB 668 TGAATATGCTAGATATGGAAGACATGATGGCGAAGCGCTCTCGGGTAACTACT 727
QY 721 GACGCTCATGACGAAAGCGTGGGAGACAAACGAGTTAGATCCCTGGTAGTCCAGCGC 780
DB 728 GACGCTCATGACGAAAGCGTGGGAGACAAACGAGTTAGATCCCTGGTAGTCCAGCGC 787
QY 781 CTAAACGATGTCAACTAGTTGTTGG--GCCTTAATTAAGCTTGGTAAAGCAAGCTAACGCGT 838
DB 788 CTAAACGATGTCAACTAGTGTGGAGGGTTAAACCTCTTAGTGTGATAGCTAACGCGT 847
```

```
QY 839 GAAGTTGACCGGCTGGGGAGTAGCGGTGCGAAGATTAAACTCAAGAATGACCGGGAC 898
DB 848 GAAGTTGACCGGCTGGGGAGTAGCGGTGCGAAGATTAAACTCAAGAATGACCGGGAC 907
QY 899 CCGCAAGACGGGTGATTAATGATTAATTAATGATGACGCAAGGAAACCTTACTACCC 958
DB 908 CCGCAAGACGGGTGATTAATGATTAATTAATGATGACGCAAGGAAACCTTACTACCC 967
QY 959 TTGATATGACGAATTTTCTAGAGATGATTAAGTGC---TTCCGGAACGGTAAACAGG 1015
DB 968 TTGATATGACGAATTTTCTAGAGATGATTAAGTGC---TTCCGGAACGGTAAACAGG 1027
QY 1016 TGCCTGATGCTGTGCTGACGCTCGTGTGAGATGTTGGTAAATGTCGCCAAGACG 1075
DB 1028 TGCCTGATGCTGTGCTGACGCTCGTGTGAGATGTTGGTAAATGTCGCCAAGACG 1087
QY 1076 CAACCTTGATTAATTAATGCAATCTTTGTTGGCACTTTAATGACCTGCGGTGACA 1135
DB 1088 CAACCTTGATTAATTAATGCAATCTTTAATGAGACTGCGGTGACA 1147
QY 1136 AACCGAAGAAAGTGGGATGACGTCAGTCTCAATGCGCTTAATGGGTAGGGCTTACA 1195
DB 1148 AACCGAAGAAAGTGGGATGACGTCAGTCTCAATGCGCTTAATGGGTAGGGCTTACA 1207
QY 1196 CGTAATTAATGACGCTGACAGAGGTTGCCAACCCGGAAGGGAGCTTAATCTCAGAAA 1255
DB 1208 CGTATTAATGACGCTGACAGAGGTTGCCAACCCGGAAGGGAGCTTAATCTCAGAAA 1267
QY 1256 GCGCGTGTAGTCCGATCGAGTGTGCACTCGACTCGATCGGTAAGTCCGAATCGCTAGTA 1315
DB 1268 GCGCGTGTAGTCCGATCGAGTGTGCACTCGACTCGATCGGTAAGTCCGAATCGCTAGTA 1327
QY 1316 ATCGCGATCAAGCATGTGCGGCTGAATATCTCCCGGCTCTTGAACAACGCCCGGTAC 1375
DB 1328 ATCGCGATCAAGCATGTGCGGCTGAATATCTCCCGGCTCTTGAACAACGCCCGGTAC 1387
QY 1376 ACCATGGAGTGGGTTTACCAAGAAAGAGTATGTAACCTTAAGAGAGGCGCTTGCCAC 1435
DB 1388 ACCATGGAGTGGGTTTACCAAGAAAGAGTATGTAACCTTAAGAGAGGCGCTTGCCAC 1447
QY 1436 GGTGAGATTCAATGACTGGGGTG 1457
DB 1448 GGTGAGATTCAATGACTGGGGTG 1469

RESULT 13
ABQ78660
ID ABQ78660 standard; DNA; 1495 BP.
XX
AC ABQ78660;
XX
DT 25-NOV-2002 (first entry)
XX
DE Nucleotide sequence of a fragment of 16S rDNA gene.
XX
KW 16S rDNA; nitrogen oxide; nitrogen; ss.
XX
OS Denitromonas aromaticus.
XX
PN JP2002142755-A.
XX
PD 21-MAY-2002.
XX
PF 14-NOV-2000; 2000JP-00346844.
XX
PR 14-NOV-2000; 2000JP-00346844.
XX
PA (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIUTSU SO.
XX
PA (BIOI-) BIOINDUSTRY KYOKAI SH.
XX
PA (AJIN) AJINOMOTO KK.
XX
DR WPI; 2002-611517/66.
XX
```

PT Microorganism belonging to Denitromonas for removal of nitrogen oxide by  
 PT reduction of nitrogen oxide to nitrogen in gas or aqueous solution, has  
 PT specific mycological characters.

XX Disclosure; Page 6-7, 8pp; Japanese.

CC The present sequence represents a fragment of 16S rDNA gene of  
 CC Denitromonas aromatics. This is a novel microorganism belonging to the  
 CC new genus Denitromonas, and which has specific mycological characters.  
 CC The new microorganism efficiently reduces and removes nitrogen oxide. The  
 CC bacterium is used for the removal of nitrogen oxide. It reduces nitrogen  
 CC oxide to nitrogen in a gas or aqueous solution

XX Sequence 1495 BP; 363 A; 350 C; 483 G; 299 T; 0 U; 0 Other;

Query Match 82.2%; Score 1197.2; DB 6; Length 1495;  
 Best Local Similarity 89.7%; Pred. No. 0; Mismatches 148; Indels 3; Gaps 2;  
 Matches 1309; Conservative 0;

1 ATTTGAACGCTGGCGCGATGCTTTACATGCAAGTGAACGCGGACGCGATGCTTGCAT 60  
 16 ATTTGAACGCTGGCGCGATGCTTTACATGCAAGTGAACGCGGACGCGATGCTTGCAT 75  
 61 CTGGTGGCAGTGGCGGACGCGGATGAGTAATGATCGGAACTGTAACGAGGCGGATA 120  
 76 CTGGTGGCAGTGGCGGACGCGGATGAGTAATGATCGGAACTGTAACGAGGCGGATA 135  
 121 AGCGATGAAAGATGCTTAATCCGCTTAATCTTAAGAGAAAGCGGCGGATGAA 180  
 136 GCCCGCGCAAGCGGATTAATCCGCTTAATCTTAAGAGAAAGCGGCGGATGAA 195  
 181 GACCTTGGCGCTTTGAGAGCGCGGATGCTTAAGTGAAGTGGTGGGATTAAGGCTTAC 240  
 196 GGGCTGGCGCTTACGAGCGGCTGATGTCGATTAAGTGAAGTGGTGGGATTAAGGCTTAC 255  
 241 CAAGCGACGATGATGATGCTTCTGAGAGACGACGACGACGACGACGACGACGACG 300  
 256 CAAGCGACGATGATGATGCTTCTGAGAGACGACGACGACGACGACGACGACGACG 315  
 301 GCCGACGCTTCTGAGAGACGACGACGACGACGACGACGACGACGACGACGACGACG 360  
 316 GCCGACGCTTCTGAGAGACGACGACGACGACGACGACGACGACGACGACGACGACG 375  
 361 CAGCATGCGCGGATGAGAAAGGCTTCCGCTTGAAGCTCTTCACTGAGAGAGA 420  
 376 CAGCATGCGCGGATGAGAAAGGCTTCCGCTTGAAGCTCTTCACTGAGAGAGA 435  
 421 AAAGTTACCGTAAATATCGTGAATCGATGACGCTATGACAGAAAGACCGGCTTAC 480  
 436 AATCATCGCCCGAATACGCGCGGATGATGACGCTATGACAGAAAGACCGGCTTAC 495  
 481 TAGGTGCGACGACCGCGGATTAATCGTGGTGCAGACGCTTAATCGAATTAATCTGGGCGT 540  
 496 TAGGTGCGACGACCGCGGATTAATCGTGGTGCAGACGCTTAATCGAATTAATCTGGGCGT 555  
 541 AAAGGTCGCGACGCGGCTTGTGAATCGATGAGTAATCGCGGCTTAATCGGGAAT 600  
 556 AAAGGTCGCGACGCGGCTTGTGAATCGATGAGTAATCGCGGCTTAATCGGGAAT 615  
 601 TGGCTTTGAATCTCAAGGCTGAGTGGCAGAGAGGAGTGAATTCATGTGTACAG 660  
 616 TGGCTTTGAATCTCAAGGCTGAGTGGCAGAGAGGAGTGAATTCATGTGTACAG 675  
 661 TGAATGCGTGAATGAGTAATGAAAGACATCGATGCGGAAAGGACGCTTGGGCTTAACT 720  
 676 TGAATGCGTGAATGAGTAATGAAAGACATCGATGCGGAAAGGACGCTTGGGCTTAACT 735  
 721 GAGCTCATGACGAAAGCGTGGGAGAGAAACAGATTAATGATTAATCGTGAATTCAGGCC 780  
 736 GAGCTCATGACGAAAGCGTGGGAGAGAAACAGATTAATGATTAATCGTGAATTCAGGCC 795  
 781 CTAAACGATGCACTAGTGTGTT--GGGCTTATTAGGCTTGTACGAACTAACGCGT 838

Db 796 CTAAACGATGCACTAGTGTGTCGAGAGAACTTTCTGGTAAACGACGCTAACGCGT 855  
 Qy 839 GAAGTTGACCGGCTGGGAGATACGCTGCGCAAGATTAAACCTCAAGGAATTGACGGGAC 898  
 Db 856 GAAGTTGACCGGCTGGGAGATACGCTGCGCAAGATTAAACCTCAAGGAATTGACGGGAC 915  
 Qy 899 CCGGACGAGCGGTGATTAATGATTAATTCGATGCAACCGCAAAACCTTACCTACCC 958  
 Db 916 CCGGACGAGCGGTGATTAATGATTAATTCGATGCAACCGCAAAACCTTACCTACCC 975  
 Qy 959 TTGACATGATGCAATTTTCTAGAGATGATTAATG-CTTGGGAAACCTTAACACAGTGG 1017  
 Db 976 TTGACATGATGCAATTTTCTAGAGATGATTAATG-CTTGGGAAACCTTAACACAGTGG 1035  
 Qy 1018 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077  
 Db 1036 CTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1095  
 Qy 1078 ACCCTTTCATTAATTTGCGATGATTTGGTGGGCACTTAAATGAGACTGCGGTGACAA 1137  
 Db 1096 ACCCTTTCATTAATTTGCGATGATTTGGTGGGCACTTAAATGAGACTGCGGTGACAA 1155  
 Qy 1138 CCGGAGAAAGTGGGAGATGACGTCAGTCTTCATGAGCCCTTAATGAGGCTTCAACAG 1197  
 Db 1156 CCGGAGAAAGTGGGAGATGACGTCAGTCTTCATGAGCCCTTAATGAGGCTTCAACAG 1215  
 Qy 1198 TAAATCAATGCGGCTGACAGAGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAAGAAC 1257  
 Db 1216 TCATCAATGCTGCGGATGACAGAGGTTGCCAACCCGCGAGGAGCTTAATCTCAAGAAC 1275  
 Qy 1258 GGGTGTGATCCGATCGGATCGGATCTGCACTGCACTCCGCGAGGAGTGGGATGCTGATAT 1317  
 Db 1276 CGATGCTGATCCGATCGGATCGGATCTGCACTGCACTCCGCGAGGAGTGGGATGCTGATAT 1335  
 Qy 1318 CCGGATGACGAGTGTGCGGATGATACGTTCCGCGCTTGTGACACCGCGCTGACAC 1377  
 Db 1336 CCGGATGACGAGTGTGCGGATGATACGTTCCGCGCTTGTGACACCGCGCTGACAC 1395  
 Qy 1378 CATGGAGTGGGTTTTCACAGAGAGAGAGTGTCTAACGTTAAAGAGAGGCGCTTGCACGG 1437  
 Db 1396 CATGGAGTGGGTTTTCACAGAGAGAGAGTGTCTAACGTTAAAGAGAGGCGCTTGCACGG 1455  
 Qy 1438 TGAATTCATGACTGGGCTG 1457  
 Db 1456 TGAATTCATGACTGGGCTG 1475

RESULT 14  
 AAC86026  
 ID AAC86026 standard; cDNA; 1460 BP.  
 AC  
 XX AAC86026;  
 AC  
 XX 06-AUG-2003 (revised)  
 DT 29-AUG-2001 (first entry)  
 XX  
 DE 16S rDNA, SBR4220.  
 XX  
 KW 16S rDNA, polyphosphate accumulating organism; PAO; probe; primer;  
 KM detection; phosphorus; waste water; sludge; ss.  
 XX  
 XX Unidentified.  
 XX  
 PN WO200146459-A1.  
 XX  
 XX 28-JUN-2001.  
 PD  
 XX  
 XX 28-DEC-2000; 2000MO-AU001611.  
 PF  
 XX 23-DEC-1999; 99AU-00004867.  
 PR  
 XX  
 PA (CRW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.  
 XX

PI Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ;  
 XX  
 DR WPI: 2001-408656/43.

PT Novel oligonucleotide probe or primer useful for detecting polyphosphate  
 PT accumulating organism in a sample, comprises a sequence that is unique to  
 PT 16S rDNA of polyphosphate accumulating organisms.

XX  
 PS Claim 4; Fig 3; 54pp; English.

XX The sequences given in AAC86021-30 represent 16S rDNA sequences from  
 CC polyphosphate accumulating organisms (PAOs). Sequences which are unique  
 CC to these 16S rDNA sequences are used to create a probe or primer for  
 CC detecting the relevant organisms. The primer/probe sequences are useful  
 CC for detecting PAO cells in a sample, by treating cells in the sample to  
 CC fix cellular contents, contacting fixed cells with the primer/probe which  
 CC is labelled with a radiolabel, a reporter group or a hapten, under  
 CC conditions which allow the probe to hybridize with 16S rRNA within the  
 CC fixed cell, removing unhybridized probe from the fixed cells, and  
 CC detecting the labeled probe-RNA hybrid by fluorescence in situ  
 CC hybridization. The primer/probe sequences are useful for identifying PAOs  
 CC that are capable of biologically removing phosphorus from waste water.  
 CC Rapid assessment of the presence of a number of PAOs in a waste water  
 CC sample, can be done using the primer/ probe sequences. They allow quick  
 CC and convenient assessment of whether a sludge or waste water sample  
 CC includes PAOs and allows quantitation of PAO cells in samples. (Updated  
 CC on 06-AUG-2003 to correct OS field.)

XX Sequence 1460 BP; 359 A; 338 C; 475 G; 288 T; 0 U; 0 Other;

XX Query Match 81.9%; Score 1194; DB 4; Length 1460;

XX Best Local Similarity 89.9%; Pred. No. 0;

XX Matches 1315; Conservative 0; Mismatches 140; Indels 7; Gaps 3;

QY 1 ATTGAACGCTGGCGGCGATCTTTACATGACATGCAAGCGACGACCGGATGCTTCAT 60  
 DB 1 ATTTAAACGTGGCGGATCTTTACATGACATGCAAGCGACGACCGGCGCAAC-- 58  
 QY 61 CTGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGGAACGTAATCCAGAAAGGGGGGTA 120  
 DB 59 CTGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGGAACGTAATCCAGAAAGGGGGGTA 118  
 QY 121 AGCGATCGAAGATGCTAATACCGCATATCTCTAAGAGAGAAAGCGGGATGCA 180  
 DB 119 AGCGAGCGAAGATGCTAATACCGCATATCTCTAAGAGAGAAAGCGGGATGCA 178  
 QY 181 GACCTTGCGCTTTGGAGCGCGCATGTCTGATTAAGCTAGTTGGGGGTAAGGCTTAC 240  
 DB 179 GACCTTGCGCTTTGGAGCGCGCATGTCTGATTAAGCTAGTTGGGGGTAATGGCTTAC 238  
 QY 241 CAAGGCGAGATCAAGTATGTTGTTCTGAGAGACGACACCACTGGGATGAGACACG 300  
 DB 239 CAAGGCGAGATCCGTAAGCGGGTCTGAGAGATGATCCGCCACACTGGGATGAGACACG 298  
 QY 301 GCCCGACCTCTTACCGGAGGACGAGTGGGAAATTTTGAACAATGGGGCGCAAGCTGATC 360  
 DB 299 GCCCGACCTCTTACCGGAGGACGAGTGGGAAATTTTGAACAATGGGGCGCAAGCTGATC 358  
 QY 361 CAGCAATGCGCGTGTGAGTAAGAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGAGAG 420  
 DB 359 CAGCAATGCGCGTGTGAGTAAGAGGCTTCCGGTTGTAAAGCTCTTTCACTGAGAGAG 418  
 QY 421 AAAGGTTACGTTAATATATGATCATGACGGTATCGACAGAGAGACACCGGCTTAC 480  
 DB 419 AATGTGTTGGTTAATATACCTGATGATGACGGTATCGACAGAGAGACACCGGCTTAC 478  
 QY 481 TACGTGCCAGACCGCGGTAATACGTAGGGTCCAGCCTTAATCCGATTTACTGGGCGT 540  
 DB 479 TACGTGCCAGACCGCGGTAATACGTAGGGTCCAGCCTTAATCCGATTTACTGGGCGT 538  
 QY 541 AAAGGTTGCGAGCGGCTTTGTAAAGTCAAGATGTGAATCCCGGGCTTAACTGGGAAT 600  
 DB 539 AAAGGTTGCGAGCGGCTTTGTAAAGTCAAGATGTGAATCCCGGGCTTAACTGGGAAT 598

QY 601 TGGCTTTGAACCTAACAGCTAGAGTGTGGACGAGGAGGTGAATTCATGTGTAGCAG 660  
 DB TGCATTTAGAGCTGACAGACTGAGAGTTTGGACAGAGGGGGGTGAATTCATGTGTAGCAG 658  
 QY 661 TGAATGCGTGAAGATATGAGAGAACATCATGTGCGAAGGAGGCTCTCTGGTTAACT 720  
 DB 659 TGAATGCGTGAAGATATGAGAGAACATCATGTGCGAAGGAGGCTCTCTGGTTAACT 718  
 QY 721 GACGCTCATGACGAGAACCGTGGGAGCAACAGATTATGATCCCTGTGTCCAGCC 780  
 DB 719 GACGCTCATGACGAGAACCGTGGGAGCAACAGATTATGATCCCTGTGTCCAGCC 778  
 QY 781 CTAACAGATGCACTAGTGTGGS--GCTTATTAAGCTTGTAAAGCAAGCTTAACTGCGCT 838  
 DB 779 CTAACAGATGCACTAGTGTGGS--GCTTATTAAGCTTGTAAAGCAAGCTTAACTGCGCT 838  
 QY 839 GAAAGTTGACCCGCTGGGAGTACGGTTCGCAAGATTAAACTCAAGAAATTGACGGGAGC 898  
 DB 839 GAAAGTTGACCCGCTGGGAGTACGGTTCGCAAGATTAAACTCAAGAAATTGACGGGAGC 898  
 QY 899 CCGCAAGACGGGTGATTAATGTGATTATTCATGCAACGCGAATACTTACTTACC 958  
 DB 899 CCGCAAGACGGGTGATTAATGTGATTATTCATGCAACGCGAATACTTACTTACC 958  
 QY 959 TTGACATGTAGCAATTTTCTAGAGATGATTAGTCT--TGGGAAACGCTAACACACAG 1015  
 DB 959 TTGACATGTAGCAATTTTCTAGAGATGATTAGTCT--TGGGAAACGCTAACACACAG 1018  
 QY 1016 TGCTGCATGAGTGTGTGTCAGCTGTGTGAGATGTTGGTTAAGTCCGCAACGACG 1075  
 DB 1019 TGCTGCATGAGTGTGTGTCAGCTGTGTGAGATGTTGGTTAAGTCCGCAACGACG 1078  
 QY 1076 CAACCTTGTATTAATTCATCATTTTGTGGGCACTTTAAAGAACTCCGGTGCA 1135  
 DB 1079 CAACCTTGTATTAATTCATCATTTTGTGGGCACTTTAAAGAACTCCGGTGCA 1138  
 QY 1136 AACCGAGAAAGGTGGGATGACGTCAGTCTGCTCACTGAGGCTTAAAGGAGGCTTCA 1195  
 DB 1139 AACCGAGAAAGGTGGGATGACGTCAGTCTGCTCACTGAGGCTTAAAGGAGGCTTCA 1198  
 QY 1196 GGTATTAATATGCGCGCTTACAGAGGTTCCAAACCCGAGAGGGGAGCTTAATCTCAGAA 1255  
 DB 1199 GGTATTAATATGCGCGCTTACAGAGGTTCCAAACCCGAGAGGGGAGCTTAATCTCAGAA 1258  
 QY 1256 GCGGCTGTAGTCCGATTCGAGATCTGCACTCCGTTGAAAGTCCGAATCGCTATA 1315  
 DB 1259 GCGGCTGTAGTCCGATTCGAGATCTGCACTCCGTTGAAAGTCCGAATCGCTATA 1318  
 QY 1316 ATCCGGATTCAGCAATGTCGCGGTGAATGTTCCCGGGCTTGTATCAACCGCCGCTCAC 1375  
 DB 1319 ATCCGGATTCAGCAATGTCGCGGTGAATGTTCCCGGGCTTGTATCAACCGCCGCTCAC 1378  
 QY 1376 ACCATGGAGTGGGTTTACACAGAAAGAGTATGTTAACTGTAAGAGAGGCGCTTGCAC 1435  
 DB 1379 ACCATGGAGTGGGTTTACACAGAAAGAGTATGTTAACTGTAAGAGAGGCGCTTGCAC 1438  
 QY 1436 GGTGAATTCATGACTGGGGTG 1457  
 DB 1439 GGTGAATTCATGACTGGGGTG 1460

RESULT 15  
 ABL40355  
 ID ABL40355 standard; DNA; 1464 BP.  
 XX  
 AC ABL40355;  
 XX  
 DT 01-JUL-2002 (first entry)  
 XX  
 DE Sequence related to an oligonucleotide for the detection of *Azarcus* sp.  
 XX  
 KW Selenic acid; MA-23; ds.



OS Unidentified.  
XX WO200146459-A1.  
XX  
XX 28-JUN-2001.  
XX  
XX 28-DEC-2000; 2000WO-AU001611.  
XX  
XX 23-DEC-1999; 99AU-00004867.  
XX  
XX (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.  
XX  
XX Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ,  
XX WPI; 2001-408656/43.  
XX  
XX Novel oligonucleotide probe or primer useful for detecting polyphosphate  
XX accumulating organism in a sample, comprises a sequence that is unique to  
XX 16S rDNA of polyphosphate accumulating organisms.  
XX  
XX Claim 4; Fig 3; 5app; English.  
XX  
XX The sequences given in AAC86021-30 represent 16S rDNA sequences from  
XX polyphosphate accumulating organisms (PAOs). Sequences which are unique  
XX to these 16S rDNA sequences are used to create a probe or primer for  
XX detecting the relevant organisms. The primer/probe sequences are useful  
XX for detecting PAO cells in a sample, by treating cells in the sample to  
XX fix cellular contents, contacting fixed cells with the primer/probe which  
XX is labeled with a radiolabel, a reporter group or a hapten, under  
XX conditions which allow the probe to hybridize with 16S rRNA within the  
XX fixed cell, removing unhybridized probe from the fixed cells, and  
XX detecting the labeled probe-RNA hybrid by fluorescence in situ  
XX hybridization. The primer/probe sequences are useful for identifying PAOs  
XX that are capable of biologically removing phosphorus from waste water.  
XX Rapid assessment of the presence of a number of PAOs in a waste water  
XX sample, can be done using the primer/ probe sequences. They allow quick  
XX and convenient assessment of whether a sludge or waste water sample  
XX includes PAOs and allows quantitation of PAO cells in samples. (Updated  
XX on 06-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 1459 BP; 360 A; 338 C; 472 G; 288 T; 0 U; 1 Other;  
Query Match 81.1%; Score 1181; DB 4; Length 1459;  
Best Local Similarity 89.7%; Pred. No. 0;  
Matches 1302; Conservative 1; Mismatches 141; Indels 7; Gaps 3;  
QY 12 GCGGCGATGCTTTACACATGCAAGTCGAAACGCGACGACGCGATCTTGCACTGTGGCGAG 71  
DB 11 GCGGCGATGCTTTACACATGCAAGTCGAAACGCGACGACGCGATCTTGCACTGTGGCGAG 68  
QY 72 TGGCGGACGCGGTGATGATGCAATCGGAACGTATCCAGAAAGAGGGGGGTAAAGCATTCGAAA 131  
DB 69 TGGCGGACGCGGTGATGATGCAATCGGAACGTATCCGTGAGTGGGGGATTAACCTACGAAA 128  
QY 132 GATGTGCTAATACCGCATATACCTTAAGAGAGAAAGCAGGGATCGAAAGACCTTGGCGT 191  
DB 129 GTTACGCTAATACCGCATATTTCTGTAGAGAGAAAGCAGGGATCGAAAGACCTTGGCGT 188  
QY 192 TTTGAGCGCGCCGATGCTGATTAAGTATGATGGGGGTAAAGGCTTACCAAGCGACGA 251  
DB 189 CTGGAGAGCGCGCATGTGCGATTAGCTTGTGGGGTAAAGGCTTACCAAGCGACGA 248  
QY 252 TCAAGTATGTTGTTCTGAGAGAGACGACGACACTGGGACTAGAGACAGGCCCAACTTC 311  
DB 249 TCCGTAAGCGGCTCTGAGAGATGATCCGACACTGGGACTAGAGACAGGCCCAACTTC 308  
QY 312 TACGGAGGACGACAGTGGGGGAAATTTTGAAGAGGCGGACGACGATCAGCAATGCGG 371  
DB 309 TACGGAGGACGACAGTGGGGGAAATTTTGAAGAGGCGGACGACGATCAGCAATGCGG 368  
QY 372 CGTGAAGTGAAGAGGCTTTCGGTGTGAAGCTCTTTCACTGAGAGAGAAAGTTACCG 431  
DB 369 CGTGAAGTGAAGAGGCTTTCGGTGTGAAGCTCTTTCACTGAGAGAGAAAGTTACCG 428

QY 432 TAAATTAATCGTGAAGTATGATGCAAGAGAGACGCGCTAACTAGTGGCCAGC 491  
DB 429 TTAATTAATCGTGAAGTATGATGCAAGAGAGACGCGCTAACTAGTGGCCAGC 488  
QY 492 AGCCGCGGTAAATACGTAAGGTGCAAGCGCTTAATCGGAATTAATGAGGCGTAAAGGTCGCG 551  
DB 489 AGCCGCGGTAAATACGTAAGGTGCAAGCGCTTAATCGGAATTAATGAGGCGTAAAGGTCGCG 548  
QY 552 AGCGGCGCTTTGTAAGTCAAGATGTAATCCCGGGGCTTAACCTGGGAATTCGCTTGA 611  
DB 549 AGCGGCTTTGTAAGTCAAGATGTAATCCCGGGGCTTAACCTGGGAATTCGCTTGA 608  
QY 612 CTACAGGCTGAGTGTGAGAGGAGGTGGAATTCATGTAGTACAGTGAATGCGTA 671  
DB 609 CTGCGAGGCTGAGTGTGAGAGGAGGTGGAATTCATGTAGTACAGTGAATGCGTA 668  
QY 672 GAGATATGGAAGAACATGATGAGCGGAAGGACGCTCTCGGTTAACTGACGCTCATGC 731  
DB 669 GAGATGTGAGAGAACATGATGAGCGGAAGGACGCTCTCGGTTAACTGACGCTCATGC 728  
QY 732 ACGAAAGCTGGGGAGCAACAGATTAATCCCTGTAGTCCACGCGCTTAAACGATGT 791  
DB 729 ACGAAAGCTGGGGAGCAACAGATTAATCCCTGTAGTCCACGCGCTTAAACGATGT 788  
QY 792 CAACATGTTGTTGG--GCCTAATTAAGGCTTGTAAACGAGTAAACGCGTGAAGTTGACCG 849  
DB 789 CAACATGTTGTTGGAGGCGTTAAACCTTTTATGTCCTGTAGTCCACGCGCTTAAACGATGT 848  
QY 850 CCGTGGGAGTACGCGTTCGCAAGTAAATCTCAAGAGATTAATGAGGAGACCGGCAACAGCG 909  
DB 849 CCGTGGGAGTACGCGGCGCAAGGCTTAAATCTCAAGAGATTAATGAGGAGACCGGCAACAGCG 908  
QY 910 GTGATTAATGAGTAAATTTGATGCAACGCGGAAACCTTAACCTTGAACATGTAG 969  
DB 909 GTGATTAATGAGTAAATTTGATGCAACGCGGAAACCTTAACCTTGAACATGTAG 968  
QY 970 GGAATTTTCTAGAGATGATTAATGATGCT--TCGGGAAACGCTTACACAGGTCGATGAGC 1026  
DB 969 GGAATTTTCTAGAGATGATTAATGATGCTTCCGAAAGAGCTTACACAGGTCGATGAGC 1028  
QY 1027 TGTGTCAGCTCGTGTGATGATGTTGGTTAAATGTCGCGCAAGAGCGCAACCTTGTGC 1086  
DB 1029 TGTGTCAGCTCGTGTGATGATGTTGGTTAAATGTCGCGCAAGAGCGCAACCTTGTGC 1088  
QY 1087 ATTAATTTGCTCATTTGTTGGGCACTTTAAATGAGACTGCGGTGACAAACCGGAGGAA 1146  
DB 1089 ATTAATTTGCTCATTTGTTGGGCACTTTAAATGAGACTGCGGTGACAAACCGGAGGAA 1148  
QY 1147 GGTGGGATGAGAGTCAAGTCTCTCAATGAGCCCTTAATGGGATTTCAACGTAATTAACAT 1206  
DB 1149 GGTGGGATGAGAGTCAAGTCTCTCAATGAGCCCTTAATGGGATTTCAACGTAATTAACAT 1208  
QY 1207 GCGCGGTACAGAGGTTGCCAACCAGGAGGAGGAGTAAATCTCAGAAAGCGCGTGTAG 1266  
DB 1209 GGTGGGTACAGAGGTTGCCAACCAGGAGGAGGAGGAGTAAATCTCAGAAAGCGCGTGTAG 1268  
QY 1267 TCCGATTCGAGTCTGCAACTGCACTCCGTAAATGCGGAATGCTTAATTCGGGATCA 1326  
DB 1269 TCCGATTCGAGTCTGCAACTGCACTCCGTAAATGCGGAATGCTTAATTCGGGATCA 1328  
QY 1327 GCATGTGCGCGGTGAATCGTTCCGCGGCTTGTACACACCGCCGTCAACATGCGGAGT 1386  
DB 1329 GCATGTGCGCGGTGAATCGTTCCGCGGCTTGTACACACCGCCGTCAACATGCGGAGT 1388  
QY 1387 GGGTTTCCAGAGAGAGTATGCTAAACCGTAAGAGGCGCGTTGCCAGCGTGAATTC 1446  
DB 1389 GGGTTTCCAGAGAGTATGCTAAACCGTAAGAGGCGCGTTGCCAGCGGAGTTCG 1448  
QY 1447 TGAATGCGGCTG 1457  
DB 1449 TGAATGCGGCTG 1459









Db 1165 GAGGAAGTGGGATGACGTC-AGTCTCATGGCCCTTATGGGTAGGGCTTCACACGTAA 1223  
Qy 1201 TACAAATGGCGCGTACAGAGGTTTCCMACCCGCGAGGGGAGCTTAATCTCAGAAAAGCGG 1260  
Db 1224 TACAAATGGCGCGTACAGAGGTTTCCMACCCGCGAGGGGAGCTTAATCTCAGAAAAGCGG 1283  
Qy 1261 TCGTAGTCCGGAATCGGAGTCTGCACTGCACTCCGTAAGTGGGAATCCCTAGTATCGC 1320  
Db 1284 TCGTAGTCCGGAATCGGAGTCTGCACTGCACTCCGTAAGTGGGAATCCCTAGTATCGC 1343  
Qy 1321 GGATCAGCATGTCGGGTGAATACGTTCCGGGCTTTGTACACACCGCCGTCACACAT 1380  
Db 1344 GGATCAGCATGTCGGGTGAATACGTTCCGGGCTTTGTACACACCGCCGTCACACAT 1403  
Qy 1381 GGGAGTGGGTTTCAACGAGAGGAGTATCTAACCGTAAGAGGGGCTTGGCCACGCTGA 1440  
Db 1404 GGGAGTGGTTCACCAAGAGGAGTATCTAACCGTAAGAGGGGCTTGGCCACGCTGA 1463  
Qy 1441 GATTCATGACTGGGGTG 1457  
Db 1464 GGGTCACTGACTGGGGTG 1480

RESULT 19  
AAC86022  
ID AAC86022 standard; cDNA; 1460 BP.  
XX  
AC AAC86022;  
XX  
DT 29-AUG-2001 (first entry)  
XX  
DE R. tenuis rdna #2.  
XX  
KW 16S rdna: polyphosphate accumulating organism; PAO; probe; primer;  
KM detection; phosphorus; waste water; sludge; ss.  
XX  
OS Rhodocyclus tenuis.  
XX  
PN MO200146459-A1.  
XX  
PD 28-JUN-2001.  
XX  
PF 28-DEC-2000; 2000MO-AU001611.  
XX  
PR 23-DEC-1999; 99AU-00004867.  
XX  
PA (CRCM-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.  
XX  
PI Hugenholtz P, Crocetti GR, Tyson GW, Blackall LJ;  
XX  
DR WPI; 2001-408656/43.  
XX  
PT Novel oligonucleotide probe or primer useful for detecting polyphosphate  
PT accumulating organism in a sample, comprises a sequence that is unique to  
PT 16S rdna of polyphosphate accumulating organisms.  
XX  
PS  
PS Claim 4; Fig 3; 54pp; English.  
XX  
XX The sequences given in AAC86021-30 represent 16S rdna sequences from  
CC polyphosphate accumulating organisms (PAOs). Sequences which are unique  
CC to these 16S rdna sequences are used to create a probe or primer for  
CC detecting the relevant organisms. The primer/probe sequences are useful  
CC for detecting PAO cells in a sample, by treating cells in the sample to  
CC fix cellular contents, contacting fixed cells with the primer/probe which  
CC is labelled with a radiolabel, a reporter group or a hapten, under  
CC conditions which allow the probe to hybridize with 16S RNA within the  
CC fixed cell, removing unhybridized probe from the fixed cells, and  
CC detecting the labeled probe-RNA hybrid by fluorescence in situ  
CC hybridization. The primer/probe sequences are useful for identifying PAOs  
CC that are capable of biologically removing phosphorus from waste water.  
CC Rapid assessment of the presence of a number of PAOs in a waste water  
CC sample, can be done using the primer/ probe sequences. They allow quick  
CC and convenient assessment of whether a sludge or waste water sample

CC includes PAOs and allows quantitation of PAO cells in samples  
XX  
SQ Sequence 1460 BP; 363 A; 340 C; 471 G; 204 T; 0 U; 2 Other;  
Query Match 79.4%; Score 1156.8; DB 4; Length 1460;  
Best Local Similarity 88.3%; Pred. No. 0;  
Matches 1291; Conservative 0; Mismatches 164; Indels 7; Gaps 3;  
Qy 1 ATTGAACCTGGCGGCATGCTTTTACATGCAAGTCGAACGGCAGCAGATGCTTGCAT 60  
Db 1 ATTGAACCTGGCGGCATGCTTTTACATGCAAGTCGAACGGTAAACCGGAAAAAC-- 58  
Qy 61 CTGTGGCAGATGGGCGGACGGGTGATATGATCGAATCGAATCGATCAGAGAAGGGGGAT 120  
Db 59 NTGGCAGCAGATGGGCGGAAAGGGTATGATATGATCGAATCGAATCGATCAGAGAAGGGGGAT 118  
Qy 121 ACGCATCGAAGAGTGTCTAATATCCGCATATCTCTAAGAGAGAAAGCAGGGGATTCGAA 180  
Db 119 ACGTACGAAAGTATACGCTAATACCGCATATCTGTGACAGAGAAAGCAGGGGATCTTAG 178  
Qy 181 GACCTTGGCGCTTTTGAAGCGCGCGAATGCTGATTAGCTAGTGGTGGGTAAAGGCTTAC 240  
Db 179 GACCTTGGCGCTTTTGAAGCGCGCGAATGCTGATTAGCTAGTGGTGGGTAAAGGCTTAC 238  
Qy 241 CAAGGCGACATCAATGATGTTGCTGAGAGAGCAGCAGCCACACTGGGACTGAGACAG 300  
Db 239 CAAGGCGACATCAATGATGTTGCTGAGAGAGATGATCCGCGACACTGGGACTGAGACAG 298  
Qy 301 GCCCAAGCTCTTACGGGAGGAGCAGATGGGGAAATTTTGGACATGGGCGCAAGCTGATC 360  
Db 299 GCCCAAGCTCTTACGGGAGGAGCAGATGGGGAAATTTTGGACATGGGCGCAAGCTGATC 358  
Qy 361 CAGCAATGCCGGTGTAGTAAAGAGCCCTTGGGTTTGAATCTCTTTCAGTCAGAGAAG 420  
Db 359 CAGCAATGCCGGTGTAGTAAAGAGCCCTTGGGTTTGAATCTCTTTCAGTCAGAGAAG 418  
Qy 421 AAAGGTTACGTAATTAATCGTACATGACGATCGATGACAGAGAAAGAACCCGCGTAA 480  
Db 419 AATGCACTGGCTAATATCTGCTGTGATGACGATGCCGTAACCGAAGAAACACCGGCTAAC 478  
Qy 481 TAGTGCCAGACAGCCCGGCTAATACGTAGGGTGCAGAGCTTAATCGGAATTTACTGGCGT 540  
Db 479 TAGTGCCAGACAGCCCGGCTAATACGTAGGGTGCAGAGCTTAATCGGAATTTACTGGCGT 538  
Qy 541 AAAGGTTGCGAGCGCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 539 AAAGGTTGCGAGCGCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGAT 598  
Qy 601 TGGCTTTGAAATACAAAGGCTAGTGTGAGTGGCAGAGGAGTGAATTTCCATGTAGTACG 660  
Db 599 TGGCTTTGACTGACACAGCTAGATGTTGACAGAGGAGTGAATTTCCATGTAGTACG 658  
Qy 661 TGAATGCTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 659 TGAATGCTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACGATTAATGATCCCTGTGATGTCAGGCC 780  
Db 719 GACGCTCATGACGAAAGCGTGGGAGCAAAACGATTAATGATCCCTGTGATGTCAGGCC 778  
Qy 781 CTAAAGATGCTAACTAATGTTGTTG--GGCCTTAATGAGCTTGTGTAAGAACTTAACGCGT 838  
Db 779 CTAAAGATGCTAACTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 836  
Qy 839 GAAGTTGACCGCTGGGAGTACGCTGCGAAGTTAAATCTCAAGAAATTTGACGGGAGC 898  
Db 839 GAAGTTGACCGCTGGGAGTACGCGCGAAGTTAAATCTCAAGAAATTTGACGGGAGC 896  
Qy 899 CCGCACAAGCGGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958  
Db 899 CCGCACAAGCGGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 956  
Qy 959 TTGACATGTAGCAATTTTCTAGATATGATGATGATGATGATGATGATGATGATGATGAT 1015



[illegible]

XX	AAQ26729	standard; cDNA; 1532 BP.
XX	AAQ26729;	
XX	25-JAN-1993	(first entry)
DT	16S rRNA gene from Bordetella bronchiseptica	
XX	rapid; examine; samples; test; identify; detect	
XX	Bordetella bronchiseptica.	
OS	JP04179480-A.	
PN		
XX	26-JUN-1992.	
PD		
XX	09-NOV-1990;	90JP-00304758.
PF		
XX	09-NOV-1990;	90JP-00304758.
PR		
XX		
XX		
PA	(ZENK-) ZENKOKU NOGOYO KYODO KUMITAI REN.	
PA	(NISE-) NIPPON SEIHUN KK.	

XX					
DR	WPI; 1992-263661/32.				
XX					
PT	16-Sr rRNA gene of <i>Bordetella bronchiseptica</i> - DNA fragment for probe and				
PT	the detection of the gene using them allowing examination of many samples				
PT	in short period.				
XX					
PS	Claim 1; Page 1; 11pp; English.				
XX					
CC	This sequence represents the 16S rRNA gene from <i>Bordetella</i>				
CC	<i>bronchiseptica</i> . The 16S rRNA gene fragment was amplified by PCR and				
CC	isolated. The <i>B. bronchiseptica</i> 16S rRNA full-length gene was cloned and				
CC	the base sequence was determined. See also AAQ26730-6				
XX					
SQ	Sequence 1532 BP; 378 A; 363 C; 490 G; 301 T; 0 U; 0 Other;				
Query Match	79.1%;	Score 1152.4;	DB 2;	Length 1532;	
Best local similarity	88.3%;	Pred. No. 0;			
Matches 1289;	Conservative	0;	Mismatches 161;	Indels 10; Gaps 3;	

Query Match	Similarity	79.1%	Score 1152.4	DB 2	Length 1532
Best Local	Similarity	88.3%	Pred. No. 0		
Matches 1289	Conservative	0	Mismatches 161	Indels 10	Gaps 3
QY	1	ATTGAAAGCTGGCGGACATGCTTTACACATGCAAGTCCGAACGGCAGCACATGCTTGCAT	60		
DB	29	ATTGAAAGCTGGCGGAGTGTCTTACACATGCAAGTCCGAACGGCAGCACAGG--GCTTCCGC	86		
QY	61	CTGGTGGGAGATGGCGGAGCGGGTAGATATCATTCGGAACGATCCAGAAAGAGGGGGGTA	120		
DB	87	CTGGTGGGAGATGGCGGAGCGGGTAGATATCATTCGGAACGATCCAGAGTGGGGGGGATA	146		
QY	121	ACGCATCGAAAGATGTCCTAATACCGCATATACTCTAAAGAGGAAGAAGCGGGATCGAAA	180		
DB	147	ACTACGCGAAAGCGTGGCTAATACCGCATATACCGCATATACCGGAGGAAAGCGGGGAGACTTTCG	206		
QY	181	GACCTTGGCGCTTTTGGAGCGGCCCATGTCTGATTAGTATGTTGGTGGGGTTAAAGGCTTAC	240		
DB	207	GGCCCTCGACATATTGGAGCGGCCCATGTCTGATTAGTATGTTGGTGGGGTTAAAGGCTTAC	266		
QY	241	CAAGGCGAGATCACTAATGTTGGTCTGAAGAGGACGACACGACACTGGGACTGAGACAG	300		
DB	267	CAAGGCGAGATCGTACTGCTTGTGAAAGAGACGACACGACCACTGGGACTGAGACAG	326		
QY	301	GCCGAGACTCTTACGGGAGGCGACAGTGGGGAATTTTGGACATGGGCGCAGCCTGATC	360		
DB	327	GCCGAGACTCTTACGGGAGGCGACAGTGGGGAATTTTGGACATGGGCGCAGCCTGATC	386		
QY	361	CAGCATGCGCGGTGATGTAAGAAAGGCGCTCGGGTTGTAAAGCTCTTTCAGTCGAGAAAGA	420		
DB	387	CAGCATGCGCGGTGATGTAAGAAAGGCGCTCGGGTTGTAAAGCTCTTTCAGTCGAGAAAGA	446		
QY	421	AAAGGTACGCTAATTAATTCGTGACTGACGTGATCGACAGAAAGACCCGGCTAAC	480		
DB	447	AAAGGTACGCGCTAATTAATTCGTGACTGACGTGATCGACAGAAAGACCCGGCTAAC	506		
QY	481	TACGTGCGACAGCGCGCGGTAAATACGTAGGGTTCGAAGCTTAAATTCGAGGCGGT	540		
DB	507	TACGTGCGACAGCGCGCGGTAAATACGTAGGGTTCGAAGCTTAAATTCGAGGCGGT	566		
QY	541	AAAGGGTGGCGAGCGCGCTTTTAAGTCAGATGTAATCCCGGGCTTAACTGGGAAT	600		
DB	567	AAAGGGTGGCGAGCGCGCTTTGAAAGTAATGTAATCCCGGGCTTAACTGGGAAT	626		
QY	601	TGCGTTTGAATTAACAGGCTGAGTGTGGCGAGAGGAGGTGGAATTCATGTGTAGCAG	660		
DB	627	TGCAATTTTAATTAACAGGCTGAGTGTGTGAGAGGAGGTGGAATTCGCGGTGTAGCAG	686		
QY	661	TGAATGCGTAGAATATGGAAGAAACATGATGGCGAAGCGCTCTGGGTTAACACT	720		
DB	687	TGAATGCGTAGAATATGGAAGAAACATGATGGCGAAGCGCTCTGGGTTAACACT	746		
QY	721	GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGTAGTCCAGCC	780		
DB	747	GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGTAGTCCAGCC	806		
QY	781	CTAAAGATGTCAACTAGTGTGGGCTTATTAAGCTTGTGAACGAACCTAAACGCGTGA	840		

Db	807	CTAAACGATGTCATTAAGCTGTGGGGCCCTTCGGGACCTTGGTAGCGCAAGCTTAAGCGCTGGA	866
Qy	841	AGTTACACCGCCCTGGGGAGTACGGTTCGCAAGATTAAATCTCAAGGAATTGAACGGGAGACC	900
Db	867	AGTTACCGCCCTGGGGAGTACGGTTCGCAAGATTAAATCTCAAGGAATTGAACGGGAGACC	926
Qy	901	GCACAAGCGGTGGATTATATGTGGATTAAATTCGATGCACGCGGAAAACTTAACTTACCTT	960
Db	927	GCACAAGCGGTGGATTATATGTGGATTAAATTCGATGCACGCGGAAAACTTAACTTACCTT	986
Qy	961	GACATGTAGCCGAATTTTCTAAGATAGATTAGTGTCT---TGGGAAACCTTAACAAGGTG	1017
Db	987	GACATGTGTGGAATCCCGAAGAGATTTGGGAGTGTCCGCAAGAGAACCGGAACACAGGTG	1048
Qy	1018	CTGCATGGCTGTGCGTCAGCTCGTGTGCGTAGATGTTGGGTTTAAGTCCCGCAACGACGCA	1077
Db	1047	CTGCATGGCTGTGCGTCAGCTCGTGTGCGTAGATGTTGGGTTTAAGTCCCGCAACGACGCA	1106
Qy	1078	ACCCCTTGTCATTTAATTTGCCATCATTTGGTGGGCACTTTAATGAGACTTCCGGTGACAA	1137
Db	1107	ACCCCTTGTCATTTAATTTGCTACGAA-----GGGCACTCTAATGAGACTTCCGGTGACAA	1161
Qy	1138	CCGGAGGAAGTGGGGATGACGTCAAGTCTTCATGCGCCCTTAATGGGTAGGGCTTCACACG	1197
Db	1162	CCGGAGGAAGTGGGGATGACGTCAAGTCTTCATGCGCCCTTAATGGGTAGGGCTTCACACG	1221
Qy	1198	TAAATCAATGCGGCGCTTACAGAGGGTTGGCCAAACCCGAGAGGGAGCTAATCTCAGAAAGC	1257
Db	1222	TCATACAAATGCTCGGGACAGAGAGGTCTGCCAACCCCGAGAGGGAGCCAAATCCAGAAAC	1281
Qy	1258	GCGTGTGATGCCGATCGGAAGTCTTGCAACTCGACTCTCGTGAAGTGGAAATGCTATGTAAT	1317
Db	1282	CGATCGTATGCCGATCGGAAGTCTTGCAACTCGACTCGTGAAGTGGAAATGCTATGTAAT	1341
Qy	1318	CGCGGATCAGCATGTCTCGCGGTGAATAAGTTCCTCCGGGTCTTGTAGCACACCGCCCGTCAAC	1377
Db	1342	CGCGGATCAGCATGTCTCGCGGTGAATAAGTTCCTCCGGGTCTTGTAGCACACCGCCCGTCAAC	1401
Qy	1378	CATGGGAGTGGGTTTCAACAGAGGAGGTAGTCTAACCGTAAAGAGGGCGCTTGCCACGG	1437
Db	1402	CATGGGAGTGGGTTTACCAAGAGTATGTTAGCTTAACCGCAAGGGGGCGATTACACGG	1461
Qy	1438	TGAGATTTCATGACTGGGGTG 1457	
Db	1462	TAGGATTTCATGACTGGGGTG 1481	
RESULT 22			
ADB61686			
ID	ADB61686	standard, DNA; 1532 BP.	
AC	ADB61686;		
DT	04-DEC-2003	(first entry)	
XX	16S rRNA of Bordetella bronchiseptica DNA sequence.		
XX	enriching mRNA; high quality bacterial mRNA; bacterial gene expression;		
XX	poly-A tail; mRNA purification; oligo-dT capture;		
XX	prokaryotic mRNA purification; bridging oligonucleotide; targeting region;		
XX	capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;		
XX	eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;		
XX	28S eukaryotic RNA bridging oligonucleotide; ds.		
XX	Bordetella bronchiseptica.		
XX	MO2003054162-A2.		
XX	03-JUL-2003.		
XX	19-DEC-2002; 2002WO-US041014.		
XX			

PR		20-DEC-2001; 2001US-00029397.	
XX			
PA	(AMBI-) AMBION INC.		
PI	Murphy GL, Whitley JP;		
XX			
DR	WPI, 2003-663255/62.		
PT	Depleting or isolating targeted nucleic acids e.g. rRNA, involves using a		
PT	bridging oligonucleotide comprising bridging region and a targeting		
PT	region complementary to a targeted nucleic acid; and a capture		
PT	oligonucleotide.		
PS	Claim 4; Page 171-172; 208pp; English.		
XX			
CC	This invention relates to a novel method for isolating, depleting or		
CC	separating a targeted nucleic acid, such as RNA, from a sample		
CC	comprising targeted and non-targeted nucleic acids. It effects a way of		
CC	enriching for non-targeted nucleic acids such as mRNAs. Isolating		
CC	sufficient quantities of high quality bacterial mRNA is a demanding		
CC	process which impedes analysis of bacterial gene expression in the		
CC	presence of host cells. A small percentage of bacterial mRNAs may be poly		
CC	-A tailed, but these are targeted for degradation and tend to be		
CC	unstable. As a result, the commonly employed method for mRNA purification		
CC	with eukaryotic cells, oligo-dT capture, is ineffective. The present		
CC	invention provides an alternative, more suitable method for mRNA		
CC	purification from prokaryotes. The method of the invention comprises the		
CC	incubation of a sample with a bridging oligonucleotide (containing a		
CC	targeting region) and subsequently incubating with a capture		
CC	oligonucleotide allowing the isolation of the target from the sample. The		
CC	method is useful for depleting or isolating targeted nucleic acid <sup>1</sup> for		
CC	example rRNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S		
CC	or 18S or eukaryotic 28S rRNA, from a sample. The rRNA sequence may		
CC	comprise any one of 64 fully defined sequences as given in the		
CC	specification. The present sequence is that of a DNA sequence which		
CC	represents the sequence of 16S rRNA of Bordetella bronchiseptica related		
CC	to the invention.		
XX			
SQ	Sequence 1532 BP; 378 A; 363 C; 490 G; 301 T; 0 U; 0 Other;		
Query Match	79.1%; Score 1152.4; DB 10; Length 1532;		
Best Local Similarity	88.3%; Pred. No. 0;		
Matches 1289; Conservative	0; Mismatches 161; Indels 10; Gaps 3		
OY	1 ATTGAACCTGGCGGCAATGCTTTACATGCAGTAGTCGACACGCGATGCTTGAT 60		
DB	29 ATGTAAAGCCTGGCGGATGCTTTACATGCAGTAGTCGACACGCG--GCTTCGCC 86		
OY	61 CTGTGCGCATGTCGCGACGCGGTGATGATGCATGGAACGTATCCAGAAGAGGGGGTAT 120		
DB	87 CTGTGTGGGAGTGGCGGAACGGGTGATGATGATGTAATCGAACCGCCAGTAAAGCGGGGATA 146		
OY	121 ACCGATCGAAAGATGTGCTAATACCGCATATCTCTAAGGAAGGAAGCGGGGATCGAAA 180		
DB	147 ACTACGCAAAAGCGTGCTAATACCGCATACGCCCTTACGGGGGAAGCGGGGACCTTCG 206		
OY	181 GACTTTGGCTTTTGGAGCGGCGCGATGTCGTATTAGTACTGATGGGTAAAGGCTTAC 240		
DB	207 GGCTTCGACATTTGGACGGCGCATTCGAGTTAGCTAAGTTGGTGGGGTAAACGGCGTAC 266		
OY	241 CAAGGCGACGATCAGTAGTTGGTCTGAGAGGACGACGACCACACTGGGACTGGAACACG 300		
DB	267 CAAGGCGACGATCCGTACCTGTTTGAAGAGGACGACGACCACACTGGGACTGGAACACG 326		
OY	301 GCCGCAATCTCTACGGGAGGACGACGATGGGGAAATTTGGACAATGGGGCGAAGCGCTGATC 360		
DB	327 GCCGCAATCTCTACGGGAGGACGACGATGGGGAAATTTGGACAATGGGGCGAAGCGCTGATC 386		
OY	361 CAGCAATCCCGTGAGTGAAGAAGAGCGCTTGGGTTGTAAAGCTCTTTCAGTCGAGAAGA 420		
DB	387 CAGCAATCCCGGTGTGCGATGAAGAGCGCTTGGGTTGTAAAGCACTTTTGGAGAGGAAGA 446		
OY	421 AAAGGTAAACGTTAAATTCGTGACTATGACGCTATGACAGAAAGAACCGGCTAAC 480		

```

Db      447 AACGCGACCGGGCTAATATCTCTGCAACTGACGGTACCTGCGAATTAACACCGGCTAAC 506
Qy      481 TAGCGCCAGCAGCCGCGGTAATAGTAGAGGTGCAAGCGTTATCGGAATTACTGGGCGT 540
Db      507 TAGCGCCAGCAGCCGCGGTAATAGTAGAGGTGCAAGCGTTATCGGAATTACTGGGCGT 566
Qy      541 AAAGGTGCGCAGCGCGCTTTGTAGTCAGATGTGAATCCCGGGCTTAACTGGGAAT 600
Db      567 AAAGCTGCGCAGCGCGCTTTCGAAAGAAAGATGTGAATCCCGGGCTTAACTGGGAAC 626
Qy      601 TCGGTTGAAACTCAAGGCTAGAGTGTGCGACAGGAGGTGGAATTCATGTGTACAG 660
Db      627 TGCATTTTAACTACCGGGCTAGAGTGTGTCAGAGGAGGTGGAATTCGCGGTGACAG 686
Qy      661 TGAATGCGTAGAGATATGGAAGACATGATGCGGAGGACGCTCTCGGTTAACTACT 720
Db      687 TGAATGCGTAGAGATATGCGGAGACACGATGCGAGGACGCTCTCGGTTAACTACT 746
Qy      721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAAGATTAAGATACCTGTAGTCAAGCC 780
Db      747 GACGCTCATGCAAGAAAGCGTGGGAGCAAAAGATTAAGATACCTGTAGTCAAGCC 806
Qy      781 CTAAACGATGTCAACTAGTTGTTGGGCTTTATTAGGCTTGTAAACGAGCTAACGGGTGA 840
Db      807 CTAAACGATGTCAACTAGTTGTTGGGCTTTATTAGGCTTGTAAACGAGCTAACGGGTGA 866
Qy      841 AGTTGACCGGCTTGGGAGATACGGTCCGAGAAATTAACTCAAGGAATTGACGGGAGCC 900
Db      867 AGTTGACCGGCTTGGGAGATACGGTCCGAGAAATTAACTCAAGGAATTGACGGGAGCC 926
Qy      901 GCAACAGCGGTGATTAATGTGATTAATGATGCAACGCGAAACCTTACCTACCTT 960
Db      927 GCAACAGCGGTGATTAATGTGATTAATGATGCAACGCGAAACCTTACCTACCTT 986
Qy      961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTCT--TCGGAGACGCTAACACAGGTG 1017
Db      987 GACATGTCTGGAAATCCCGAAGAGATTTGGAGTGTCTGCAAGAGAACCGGAAACACAGGTG 1046
Qy      1018 CTGCAATGAGCTGTCTGACGCTCTGCTCTGATGCTTGGGTTAAGTCCCGAAGAGGCGCA 1077
Db      1047 CTGCAATGAGCTGTCTGACGCTCTGCTCTGATGCTTGGGTTAAGTCCCGAAGAGGCGCA 1106
Qy      1078 ACCCTTGTCAATTAATGCAATCATTTGTTGGGCACTTAAATGAGACTGCGCGGTGACAA 1137
Db      1107 ACCCTTGTCAATTAATGCTACGAAA-----GGGCACTTAATGAGACTGCGCGGTGACAA 1161
Qy      1138 CCGAGAGAGGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTCAACAG 1197
Db      1162 CCGAGAGAGGTGGGAGTGAAGTCAAGTCTCATGAGCCCTTATGGGTAGGGCTTCAACAG 1221
Qy      1198 TAATCAATGCGCGGTACAGAGGTTGCGAACCCCGGAGGGGGAGCTTAATCTGAAAGC 1257
Db      1222 TCATACATATGTTGCGGACAGAGGTTGCGAACCCCGGAGGGGGAGCCATCTCCGAAACC 1281
Qy      1258 GCGTCGATAGTCGAGATCGAGTCTGCAACTGCACTCCGTGAAGTCGGAATCGTAGTAT 1317
Db      1282 CGATCGATAGTCGAGATCGCACTCTGCAACTGCACTGCGTGAAGTCGGAATCGTAGTAT 1341
Qy      1318 CGCGGATCGAGATGTCGCGGTGAATACGTTCCGGGCTTTGTACACACCGCCGCTCACAC 1377
Db      1342 CGCGGATCGAGATGTCGCGGTGAATACGTTCCGGGCTTTGTACACACCGCCGCTCACAC 1401
Qy      1378 CATGGAATGGGTTTACCAAGACAGAGTACTTAACCTTAAGAGAGGGCGGTTTGCACGG 1437
Db      1402 CATGGAATGGGTTTACCAAGAGTACTTAACCTTAACCGCAAGGGGGCGGATTAACACGG 1461
Qy      1438 TGAGATTGATGACTGGGGTG 1457
Db      1462 TAGGATTCATGACTGGGGTG 1481

```

RESULT 23

```

AA01866
ID      AA01866 standard; DNA; 1526 BP.
XX
AC      AA01866;
XX
DT      16-OCT-2003 (revised)
DT      03-AUG-1999 (first entry)
XX
DE      P. cepacia 16S rRNA gene sequence.
XX
KW      16S rRNA; KK01; primer; PCR; amplification; probe; hybridisation;
KW      detection; diagnosis; de.
XX
OS      Burkholderia cepacia.
XX
PN      JF07255486-A.
XX
PD      09-OCT-1995.
XX
PE      23-MAR-1994; 94UP-00051739.
XX
PR      23-MAR-1994; 94UP-00051739.
XX
PA      (CANO ) CANON KK.
XX
DR      WPI; 1995-378541/49.
XX
PT      Pseudomonas cepacia KK01 strain 16S rRNA gene - also related probes and
PT      primers, useful for specific detection of P. cepacia strain KK01.
XX
PS      Claim 1; Page 21; 21pp; Japanese.
XX
CC      This sequence represents the 16S rRNA gene of Pseudomonas cepacia strain
CC      KK01. Fragments of the nucleic acid sequence (see AA01872-T02316) are
CC      useful as primers and probes for the specific detection of P. cepacia
CC      strain KK01. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ      Sequence 1526 BP; 382 A; 352 C; 486 G; 306 T; 0 U; 0 Other;
Query Match      78.8%; Score 1148.8; DB 2; Length 1526;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 172; Indels 8; Gaps 2;
Qy      1 ATTGAACGCTGCGCGGATGCTTTTACACATGCAAGTGAACCGGACGAGACGCGATGTCAT 60
Db      21 ATTGAACGCTGCGCGGATGCTTTTACACATGCAAGTGAACCGGACGAGACGCGATGTCAC 80
Qy      61 CTGTCGCGAGTGGCGGAGCGGATGAGTAATGATCGGAACGTAATCCAGAAAGAGGGGATA 120
Db      81 CTGTCGCGAGTGGCGGAGCGGATGAGTAATGATCGGAACGTAATCCAGAAAGAGGGGATA 140
Qy      121 ACCGATCGAAAGATGTCGTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGGATCGAAA 180
Db      141 GCCCGGGAAGAGCGGATTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGGATCGAAA 200
Qy      181 GACCTTGGCTTTTGGAGCGCGGATGTCGATTAAGTGAATGTTGGTGGGCTTAAGGCTTAC 240
Db      201 GGCCTGCGCTTAATAGGCTTGGCGATGCTGATTAAGTGAATGTTGGTGGGCTTAC 260
Qy      241 CAAGCGGACGATCAGTATGTTGTTGAGAGAGACAGACGACCACTGGGACTGAGACAGC 300
Db      261 CAAGCGGACGATCAGTATGTTGTTGAGAGAGACAGACGACCACTGGGACTGAGACAGC 320
Qy      301 GCCGAGACTCTTAACGGGAGCAGCAGTGGGAAATTTTGAACATGGGCGACGACTGATC 360
Db      321 GCCGAGACTCTTAACGGGAGCAGCAGTGGGAAATTTTGAACATGGGCGGAAAGCTGATC 380
Qy      361 CAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGAAGCTTTTCAAGTGAAGAGA 420
Db      381 CAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGAAGCACTTTTGTCCGAAAGA 440
Qy      421 AAAGTTACGTTAATATGTCATCATGACGATGACAGAGGATCGACAGAAAGACGCGGCTAAC 480

```

```
Db 441 AATCGTGGCTGTAATACAGTCGGGGGATGACGGTACCGGAAGTAATACACCGGCTAAC 500
Qy 481 TACGTGCCAGACCGCGGGTAACTAGCTAGGTCGCAAGCCTTAATCGGAATTACTGGGCT 540
Db 501 TACGTGCCAGACCGCGGGTAACTAGCTAGGTCGCAAGCCTTAATCGGAATTACTGGGCT 560
Qy 541 AAAAGGTGGCAGCGGCTTTTGTGAATGTCAGATGTGAAATCCCGGGCTTAACCTGGGAAT 600
Db 561 AAAAGGTGGCAGCGGCTTTTGTGAATGTCAGATGTGAAATCCCGGGCTTAACCTGGGAAT 620
Qy 601 TCCGTTTGAACCTACAAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTAGCAG 660
Db 621 TCCATTTGGTGACTGGCAAGCTAGATGTGGCAGAGGAGGAGGTGGAATTCACGCTGTAGCAG 680
Qy 661 TGAATGCGTACAGATATGGAAGAACATGTGATGGCGAAGCGCTCTCTGGGTAACT 720
Db 681 TGAATGCGTACAGATATGGAAGAACATGTGATGGCGAAGCGCTCTCTGGGTAACT 740
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGTACCCTGGTGTCCAGCC 780
Db 741 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGTACCCTGGTGTCCAGCC 800
Qy 781 CTAAACGATGTCACTAGTGTGTGGGCTTATTAAGCTTGTGAACGAGCTTAACGGGTGA 840
Db 801 CTAAACGATGTCACTAGTGTGTGGGCTTATTAAGCTTGTGAACGAGCTTAACGGGTGA 860
Qy 841 AGTTGACCGCTGGGAGTACCGTCCGAAAGATTTAAACTCAAAGGAATTGACGGGAGCC 900
Db 861 AGTTGACCGCTGGGAGTACCGTCCGAAAGATTTAAACTCAAAGGAATTGACGGGAGCC 920
Qy 901 GGACAGAGCGGTGATATGATGATTAATTCGATGCAACCGGAAAACTTACCTACCTT 960
Db 921 GGACAGAGCGGTGATATGATGATTAATTCGATGCAACCGGAAAACTTACCTACCTT 980
Qy 961 GACATGTACGGAATTTTCTAGATATGATTAAGTGTCT--TCGGGAACGCTAACACAGGTG 1017
Db 981 GACATGTGTGGAATCCTGCTGAGAGGTGGAGTGTCTGAAAGGAACCGGCGCACAGGTG 1040
Qy 1018 CTGCAATGGTGTGCTGCACTGCTGTGTGTGAATGTTGGGTAAAGTCCCGCAACGAGCGCA 1077
Db 1041 CTGCAATGGTGTGCTGCACTGCTGTGTGTGAATGTTGGGTAAAGTCCCGCAACGAGCGCA 1100
Qy 1078 ACCCTGTGCTTAATTAATGCAATCAATTTGTTGGGCACTTAAATGAGACTCCGGTGCACA 1137
Db 1101 ACCCTGTGCTTAATTAATGCAATCAATTTGTTGGGCACTTAAATGAGACTCCGGTGCACA 1155
Qy 1138 CCGAGAGGAAGTGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTCAACG 1197
Db 1156 CCGAGAGGAAGTGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTCAACG 1215
Qy 1198 TAAATCAATGCGCGCTACAGAGGGTTGCCAACCCCGGAGGGGAGCTTAATTCAGAAAGC 1257
Db 1216 TCAATCAATGCTCGGAACAGAGGGTTGCCAACCCCGGAGGGGAGCTTAATTCAGAAAGC 1275
Qy 1258 GCGTGTGTGTCGGGAATCGGAAGTCTGCAACTCGGACTCGGGAATGGGAATGCTGTATAT 1317
Db 1276 CGATGTGTGTCGGGAATCGGAAGTCTGCAACTCGGACTCGGGAATGGGAATGCTGTATAT 1335
Qy 1318 CCGGATACGATGTCGCGGTAAATAGTTCCTCGGGTCTTGTATACACCGCCCGTCAAC 1377
Db 1336 CCGGATACGATGTCGCGGTAAATAGTTCCTCGGGTCTTGTATACACCGCCCGTCAAC 1395
Qy 1378 CATGGAGTGGGTTTCAACGAAGCAGGTAGTCTAAACGTTAAGAGAGGGCGCTTGCACAG 1437
Db 1396 CATGGAGTGGGTTTCAACGAAGCAGGTAGTCTAAACGTTAAGAGAGGGCGCTTGCACAG 1455
Qy 1438 TGAAGATTCACTGAGCTGGGCTG 1457
Db 1456 TGAAGATTCACTGAGCTGGGCTG 1475
```

RESULT 24  
AD016355

```
ID AD016355 standard; DNA; 1496 BP.
XX
AC AD016355;
XX
DT 09-SEP-2004 (first entry)
XX
DE Nucleotide sequence of a Bordetella variable 16S rRNA gene region.
XX
KW fragmentation-based method; mass spectrometric method;
KW nucleic acid polymorphism; nucleic acid mutation; genetic disease;
KW chromosome abnormality; obesity; atherosclerosis; cancer; haplotype; ss;
KW rRNA gene.
XX
OS Bordetella pertussis.
XX
PN W02004050839-A2.
XX
PD 17-JUN-2004.
XX
PF 26-NOV-2003; 2003MO-US037931.
XX
PR 27-NOV-2002; 2002US-0429895P.
XX
PA (SEQU-) SEQUENOM INC.
XX
PI Van Den Boom D, Boecker S;
XX
DR WPI; 2004-487567/46.
XX
PT Use of fragmentation-based methods and systems, e.g. mass spectrometric
PT methods for the analysis of sequence variations including nucleic acid
PT polymorphisms and mutations.
XX
PS Example 4; SEQ ID NO 32; 198pp; English.
XX
CC The specification describes the use of fragmentation-based methods and
CC systems including mass spectrometric methods for the analysis of sequence
CC variations including nucleic acid polymorphisms and mutations. The
CC fragmentation-based methods and systems of the invention are useful for
CC the analysis of sequence variations including nucleic acid polymorphisms
CC and mutations. The methods are useful for identifying a genetic disease
CC or chromosome abnormality; identifying a predisposition to a disease or
CC condition including obesity, atherosclerosis, or cancer; identifying an
CC infection by an infectious agent; providing information relating to
CC identity, heredity, or histocompatibility; identifying pathogens; or
CC determining haplotypes. AD016353-AD016361 represent Bordetella variable
CC 16S rRNA gene regions. Amplicons from this region were used to demonstrate
CC the invention. They were used to demonstrate a method for bacterial
CC typing by base-specific fragmentation.
XX
SQ Sequence 1496 BP; 372 A; 355 C; 475 G; 293 T; 0 U; 1 Other;
XX
Query Match 78.5%; Score 1143.8; DB 12; Length 1496;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 1281; Conservative 0; Mismatches 163; Indels 10; Gaps 3;
Qy 7 CGCTGGCGGCAATGCTTACATGCAATGCAACGCGACGAGTGTTCATCTGTTG 66
Db 1 CGCTAGCGGGATGCTTACATGCAATGCAACGCGACGCGGA--CTTGCGTGTGCG 58
Qy 67 GCGAGTGGCGGACGGGGTGAATGCAATGCAACGCGATCTATCCAGAAAGAGGGGTAAACGAT 126
Db 59 GCGAGTGGCGGACGGGGTGAATGCAATGCAACGCGATGCGCGAGTACGGGGGTAAACGAT 118
Qy 127 CGAAAGATGTCTAATACCGATATATCTTAAGAGAGAAACAGGAGATCGAAAGACCTT 186
Db 119 CGAAAGCTTACTAATACCGATATACGCTTACGCGGGGAGAAACGCGGGGACCTTGGGCTTC 178
Qy 187 GCGCTTTTGGAGCGGCGCATGTCTGATTAAGCTAGTGTGGGGGTAAAGCCTTACCAAGC 246
Db 179 GCACATATTGAGCGGCGCATATCGGATTAAGCTAGTGTGGGGGTAAAGGCTTACCAAGC 238
Qy 247 GACGATCAGTATGTTGTGTGAGAGAGACGACGACCACTGGGACTGAGACAGGCGCCAG 306
```





```

SQ      Sequence 1496 BP; 372 A; 355 C; 475 G; 293 T; 0 U; 1 Other;
Query Match      78.5%; Score 1143.8; DB 14; Length 1496;
Beet Local Similarity 88.1%; Pred. No. 0;
Matches 1281; Conservative 0; Mismatches 163; Indels 10; Gaps 3;

QY      7 CGCTGGCCGCGCATCTTTACATGCAAGTCGAACGGCCAGCAGGATGCTGCATCGGGG 66
DB      1 CGCTAGCGGAGATGCTTTTACATGCAAGTCGAACGGCCAGCAGGATGCTGCATCGGGG 58
QY      67 GCGAGTGGCGGACGGGTGAGTAATGATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 126
DB      59 GCGAGTGGCGGACGGGTGAGTAATGATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 118
QY      127 GCAAAAGATGCTTAATACCGCATATATCTTAAAGAGAAAGCAAGGGGATCGAAACCTT 186
DB      119 CAAAAGCTTAGCTTAATACCGCATATACCTTACGGGGGAAAGCGGGGACCTTCGGGCTTC 178
QY      187 GCGCTTTTGGAGCGGCGCATGCTGATTAAGTAAAGTAAAGGCTTAAGGCTTAAGGCTTAAG 246
DB      179 GCACTATTGAGAGCGGCGCATGCTGATTAAGTAAAGTAAAGGCTTAAGGCTTAAGGCTTAAG 238
QY      247 GACGATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 306
DB      239 GACGATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 298
QY      307 ACTCTACGGGAGGACGAGTGGGGAATTTTGGACATATGGGCGCAAGCTTATCCAGCA 366
DB      299 ACTCTACGGGAGGACGAGTGGGGAATTTTGGACATATGGGCGCAAGCTTATCCAGCA 358
QY      367 TGGCGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 426
DB      359 TCCCGGATGCGATGAAAGGCTTCGGGTTGTAAGCACTTTTGGCAGAAAGAAAGGCGC 418
QY      427 TACGCTAAATTAATCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 486
DB      419 TCTGCTAAATTAATCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 478
QY      487 CCAAGCAGCCGCGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
DB      479 CCAAGCAGCCGCGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538
QY      547 TGGCGGAGGCGGCTTTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 606
DB      539 TGGCGAGGCGGCTTTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 598
QY      607 TGAATCTAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 666
DB      599 TTTAATCTAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 658
QY      667 GCGTGAAGATTAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
DB      659 GCGTGAAGATTAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
QY      727 CATGACGAAAGAGTGGGAGGAGCAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 786
DB      719 CATGACGAAAGAGTGGGAGGAGCAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 778
QY      787 GATGCTCAATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 846
DB      779 GATGCTCAATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 838
QY      847 CCGCTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906
DB      839 CCGCTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 898
QY      907 GCGGAGTAAATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 966
DB      899 GCGGAGTAAATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 958
QY      967 TAGCGAATTTTCTAGAGATAGATAGTCT--TCGGGAACGCTAACAAGGTGCTGAT 1023
DB      959 TCTGGAATTCGGAAGAGATTTGGCAGTGTCTCCAGAGAGAACCGGAACAGGTGCTGAT 1018
```

```

QY      1024 GCGTGTCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1083
DB      1019 GCGTGTCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
QY      1084 GTCAATTAATGCGATCATTTGTTGGGCACTTTAATGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1143
DB      1079 GTCAATTAATGCGATCATTTGTTGGGCACTTTAATGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
QY      1144 GAAGTGGGAGTGAAGTCAATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1203
DB      1134 GAAGTGGGAGTGAAGTCAATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1193
QY      1204 AATGCGGCTACAGAGGTTTCCCAACCGCGAGGGGGAGCTAATCTCAGAAAGCGCTG 1263
DB      1194 AATGCTGGGAGCAGAGGCTCCCAACCGCGAGGGGGAGCTAATCTCAGAAAGCGCTG 1253
QY      1264 TAGTCCGATGCGAGTCTGCAACTGACTCCGTAAGTCCGTAAGTCCGTAAGTCCGTAAGTCCGTAAGTCCGTAAG 1323
DB      1254 TAGTCCGATGCGAGTCTGCAACTGACTCCGTAAGTCCGTAAGTCCGTAAGTCCGTAAGTCCGTAAGTCCGTAAG 1313
QY      1324 TCAGCATGTCGCGGTGAATACGTTCCGCGTCTTGTACACACCGCCGCTCACACATGGG 1383
DB      1314 TCAGCATGTCGCGGTGAATACGTTCCGCGTCTTGTACACACCGCCGCTCACACATGGG 1373
QY      1384 AGTGGGTTTCAACGAAGCAGTATGCTTACCGTAAGAGGGGCGCTTGGCCAGCGTGAAT 1443
DB      1374 AGTGGGTTTCAACGAAGCAGTATGCTTACCGTAAGAGGGGCGCTTGGCCAGCGTGAAT 1433
QY      1444 TCATGACTGGGGTG 1457
DB      1434 TCATGACTGGGGTG 1447
```

```

RESULT 26
AAC86029
ID      AAC86029 standard; cDNA; 1426 BP.
XX
AC      AAC86029;
XX
XX      06-AUG-2003 (revised)
DT      29-AUG-2001 (first entry)
XX
DE      16S rDNA, GCP112.
XX
XX      16S rDNA; polyphosphate accumulating organism; PAO; probe: primer;
KW      detection; phosphorus; waste water; sludge; ss.
XX
OS      unidentified.
XX
PN      WO200146459-A1.
XX
PD      28-JUN-2001.
XX
PF      28-DEC-2000; 2000WO-AU001611.
XX
PR      23-DEC-1999; 99AU-00004867.
XX
XX      (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.
XX
PI      Hugenholz P, Crocetti GR, Tyson GW, Blackall JL;
XX
XX      WPI: 2001-408656/43.
XX
DR      Novel oligonucleotide probe or primer useful for detecting polyphosphate
PT      accumulating organism in a sample, comprises a sequence that is unique to
PT      16S rDNA of polyphosphate accumulating organisms.
XX
PS      Claim 4; Fig 3; 54pp; English.
XX
CC      The sequences given in AAC86021-30 represent 16S rDNA sequences from
CC      polyphosphate accumulating organisms (PAOs). Sequences which are unique
CC      to these 16S rDNA sequences are used to create a probe or primer for
```



detecting the relevant organisms. The primer/probe sequences are useful for detecting PAO cells in a sample, by creating cells in the sample to fix cellular contents, contacting fixed cells with the primer/probe which is labelled with a radiolabel, a reporter group or a hapten, under conditions which allow the probe to hybridize with 16S rRNA within the fixed cell, removing unhybridized probe from the fixed cells, and detecting the labeled probe-RNA hybrid by fluorescence in situ hybridization. The primer/probe sequences are useful for identifying PAOs that are capable of biologically removing phosphorus from waste water. Rapid assessment of the presence of a number of PAOs in a waste water sample, can be done using the primer/ probe sequences. They allow quick and convenient assessment of whether a sludge or waste water sample includes PAOs and allows quantitation of PAO cells in samples. (Updated on 06-AUG-2003 to correct OS field.)

Sequence 1426 BP; 361 A; 335 C; 455 G; 275 T; 0 U; 0 Other;

Query Match	78.4%	Score 1142	DB 4	Length 1426
Best Local Similarity	89.3%	Pred. No. 0		
Matches 1266	Conservative	0	Mismatches 145	Indels 7
				Gaps 3

OY	12	GGGGATGCTTAAACAATCAATCAAGTGAAGGAGCAGACGGATCTTGCATCTGAGTGGCGAG	71
Db	11	GGGGATGCTTAAACAATCAATCAAGTGAAGGAGCAGACGGAGGCGAACCC--CTGTGGCAG	68
OY	72	TGGCGCAGCGGCTGAGTGAATGATCGGAACGTATCCGAAGAAGGAGGGGTAAACGCATCGAA	131
Db	69	TGGCGAAGCGGGTGAAGTAAAGATCCGAAAGTGCCTCGAATGGGGGATTAACGTACGAA	128
OY	132	GATGTGCTAATACCCGATTAATCTTAAGAGGAAGACAGGGGATCGAAGACCTTGGCGT	191
Db	129	GTTAGCGCTAATACCGCATATTCGTGAGCAGGAAGACAGGGGATCGCAAGACTTGGCGTT	188
OY	192	TTTGAAGCGGCGGATGCTGATTTAGCTGTGGTGGGGTAAAGGCGCTTACCAAGCGACGA	251
Db	189	CGAGGAACGGCGCGAGTCCGAAATTAAGTATGGTGGGGTAAAGGCTTACCAAGCGACGA	248
OY	252	TCAGTAGTTGATCTGAGAGGACGACACCGACCTGGGACTGAGACACGGGCCAAGCTCC	311
Db	249	TCCGTAGCGGGGTCTGAGAGATGATCCGCCACCTGGAATGAGACACGGTCCAGACTCC	308
OY	312	TACGGGAGGACAGTGGGGAAATTTTGGACATGGGCGCAAGCTGATCCAGCAATGGCG	371
Db	309	TACGGGAGGACACAGTGGGGAAATTTTGGACATGGGCGCAAGCTGATCCAGCAATGGCG	368
OY	372	CGTAGTGAAGAAGCCTCGGGTTGAAAGCTCTTCAAGTCGAGAAAGAAAGTTTACCG	431
Db	369	CGTAGTGAAGAAGCCTTCGGGTTTAAAGCTCTTTCGGCGGGAAAGAAATCGACGGG	428
OY	432	TAAATATCGTGAACCTATGACGGTATCGACAGAGAAGACACGGGCTTAATCAGTGCACG	491
Db	429	TAAATACCTGTGTGATGACGGTATCGGAAATGAAGACACGGGCTTAATCAGTGCACG	488
OY	492	AGCCGCGTAAATCGTAGGGTGCAGCGTTAATCGGAATTAATCTGGGCGTAAAGGTCGC	551
Db	489	AGCCGCGTAAATCGTAGGGTGCAGCGTTAATCGGAATTAATCTGGGCGTAAAGGTCGC	548
OY	552	AGCGGCTTTGTAAATGACAGATGGAATCCCGGGGCTTAACTGGGGAATTCGCTTTGAA	611
Db	549	AGCGGCTTTGTAAATGACAGATGGAATCCCGGGGCTCAACTGGGGAATTCGCTTTGAA	608
OY	612	CTTCAAGGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTAGCAGTGAATGCTA	671
Db	609	CTTCCAAAGCTGGAATTTGGCAGAGGGGGGTGGAATTCAGTGTAGCAGTGAATGCTA	668
OY	672	GAGATATGGAAGAACATCGATGGCGAAGCGCAGCTCTGGGTTTAACTGACGCTCATGC	731
Db	669	GAGATGTGGAGGAACACCGATGGCGAAGCGCAGCTCTGGGCTCAATTAAGCCTCATGC	728
OY	732	ACGAAAGCGGGGAGCAACAGGATTAAGTACCTGTGTAGTCCACGCGCTTAAACGATGT	791
Db	729	ACGAAAGCGGGGAGCAACAGGATTAAGTACCTGTGTAGTCCACGCGCTTAAACGATGT	788

QY	792	CAACTACTGTTGG--GCCCTATTGGCTTGGTAACGAAGCTTAACGGGTGAAGTTGACCG	849
Db	789	CAACTAGTGTGTGGAGGGTTAAACCTTTAGTGTCCCTACTTAACGGGTAAAGTTGACCG	848
QY	850	CCTGGGGAGTAACGGTCCGACAAGATTAAACTCAAGAAATTGACGGGGACCCGCAACAGCG	909
Db	849	CCTGGGGAGTACGGCCGCCCAAGGCTAAAACTCAAGAAATTGACGGGGACCCGCAACAGCG	908
QY	910	GTGAGATTATGTGATTATTAATTCGATGCAACGGGAAAACTTTACCTACCCTTGAACATGTAG	969
Db	909	GTGAGATATGTGATTATTAATTCGATGCAACGGGAAAACTTTACCTACCCTTGAACATGTCA	968
QY	970	CGAATTTTCTTAGAATATAGATTAGTGC---TCGGGAACGGTAAACACAGAGTGCTCATAGC	1026
Db	969	GGAATCCCGGAGAAATTTGGAGTGTCTCCAGAGAAACCTGAACACAGAGTCTCATAGC	1028
QY	1027	TGTCGTACGTCGATGTGTGAGATGTTGGTTAAAGTCCCGCAAGACGGCAACCTTTGTC	1086
Db	1029	TGTGTCAGCTCTGTCTGTGTGAGATGTTGGTTAAAGTCCCGCAAGACGGCAACCTTTGTC	1088
QY	1087	ATTATATGCCATCAATTTGGTTGGCACTTAAATAGACATGTCGGGTGACAAACCGGAGAA	1146
Db	1089	ATTATATGCCATCAATTCAGTTGGGACACTTTAATAGACTGCCGGGTGACAAACCGGAGAA	1148
QY	1147	GTGTGGGATGACGTCAAGTCCCTCAATGGCCCTTAATGGGATGGGCTTCAACGTAATTAACAT	1206
Db	1149	GGTGGGGATGACGTCAAGTCCCTCAATGGCCCTTAATGGGATGGGCTTCAACAGTCAATACAT	1208
QY	1207	GGCGCGTACAGAGGGTTGCCAACCACGGAGGGGAGCTAATCTGAAAGGCGTCTGTAG	1266
Db	1209	GGTCTGTCAAGAGGTTGCCAACCACGGAGGGGAGCCAAATCCGAAAGCCGATCTGTAG	1268
QY	1267	TCCGAGATCGAGTCTGCACACTCGACTCGTGAAGTCGGAATCGCTAGTAATTCGGGATCA	1326
Db	1269	TCCGAGATCGAGTCTGCACACTCGACTCGTGAAGTCGGAATCGCTAGTAATTCGGGATCA	1328
QY	1327	GCATGTCCGCGTGAATACGTTCCCGGGCTTTGTACACACCCGCGTCAACATCGGGAGT	1386
Db	1329	GCATGTCCGCGGTGAATACGTTCCCGGGCTTTGTACACACCCGCGTCAACATCGGGAGC	1388
QY	1387	GGGTTTCAACGAAGCAGTAGTAGTAAACCGTAAGGAGG 1424	
Db	1389	GGGTTTCTGCCAGAGTAGTAGTAAACCGCAAGGAGG 1426	
RESULT 27			
AA064008			
ID	AA064008	strand: rRNA; 1532 BP.	
XX	AA064008;		
AC			
XX	27-AUG-2003	(revised)	
DT	22-JUL-1994	(first entry)	
DT			
XX	16S rRNA gene.		
DE			
XX	16S rRNA; probe; detection; procine atrophic rhinitis; hybridisation;		
KM	Bordetella bronchiseptica; pig raising; ss.		
XX			
OS	Bordetella bronchiseptica.		
XX			
PN	JP0536999-A.		
XX			
PD	21-DEC-1993.		
XX			
PF	10-JUN-1992;	92JP-00150688.	
XX			
PR	10-JUN-1992;	92JP-00150688.	
XX			
PA	(NISE-) NIHON SEIFUN KK.		
PA	(ZENK-) ZENKOKU NOGYO KD RENGOKAI.		
XX			
DR	WPI; 1994-037379/05.		



XX (ADRE-) ADVANCED RES & TECHNOLOGY INST.  
 XX PA  
 XX PI Martin WJ, Wisniewski P;  
 XX WPI, 1998-240079/21.  
 DR Use of oligo:nucleotide(s) corresponding to bacterial 16S rRNA - for  
 PT inhibiting bacterial protein expression and treating bacterial infection.  
 XX Claim 26; Page 62-63; 73pp; English.  
 XX This polynucleotide comprises the 16S ribosomal RNA (rRNA) gene of  
 CC Burkholderia cepacia. The invention relates to methods and compositions  
 CC for the treatment of gram-negative bacterial infections employing novel  
 CC oligonucleotides as anti-microbial agents. The oligonucleotides are  
 CC targeted to the Shine-Dalgarno (SD) region of prokaryotes to inhibit  
 CC bacterial expression and hence inhibit bacterial infection. They  
 CC preferably comprise 10-35 consecutive bases of the 3' end of a bacterial  
 CC 16S rRNA (see also AAV24291-94). An oligonucleotide may also include a  
 CC transport moiety and may have DNA phosphate modifications to increase  
 CC nuclease resistance, or may be formulated in a liposome. A claimed method  
 CC for treating a bacterial infection of a patient comprises administering a  
 CC liposomal formulation of such an oligonucleotide. The oligonucleotides  
 CC can be used particularly for treating bacterial infections in pulmonary  
 CC diseases such as cystic fibrosis or tuberculosis. Since the SD sequence  
 CC is not present in eukaryotic cells, the oligonucleotides provide a  
 CC pathogen-specific therapeutic method  
 XX  
 XX Sequence 1535 BP; 385 A; 355 C; 488 G; 307 T; 0 U; 0 Other;  
 SQ  
 Query Match 78.1%; Score 1138.4; DB 2; Length 1535;  
 Best Local Similarity 87.7%; Pred. No. 0;  
 Matches 1280; Conservative 0; Mismatches 171; Indels 9; Gaps 3;

DB 568 AAGGTCGCGAGCGGGCTTGTCTAAGACCGATGTGMAATCCCGGGCTCAACTGGGAAC 627  
 QY 601 TCGCTTTGMAAATCAAGGCTAGAGTGTGGAGGAGGAGTGAATTCATGTAGAG 660  
 DB 628 TGCATTGTGTGACTGGCAGGCTAGAGTATGCGAGAGGGGGTAGAATTCACGTAGCAG 687  
 QY 661 TGAATGTGTAGAGATATGGAAGAACATCGATGGCGAAGGAGCCTCTGGGTAACT 720  
 DB 688 TGAAATGTGTAGAGATATGGAAGAACATCGATGGCGAAGGAGCCTCTGGGGCAATCT 747  
 QY 721 GACGCTCATGACGAAACGTTGGGAGGAGCAACAGATTAGATCCCTGTAGTCCACCC 780  
 DB 748 GACGCTCATGACGAAAGGTTGGGAGGAGCAACAGATTAGATCCCTGTAGTCCACCC 807  
 QY 781 CTAAACGATGTCACTAGTTGTGGGCTTATAGGCTTGTGTAAGAGCTAACCGTGA 840  
 DB 808 CTAAACGATGTCACTAGTTGTGGGAGTTCATTTCTTATAGTAACTAACCGGTGA 867  
 QY 841 AGTTGACGCGCTGGGAGAGTACGCTGCAAGATTAAACTCAAGGAAATTGACGGGACC 900  
 DB 868 AGTTGACGCGCTGGGAGAGTACGCTGCAAGATTAAACTCAAGGAAATTGACGGGACC 927  
 QY 901 GCACAAGCGGTGATATGTGATTAATTGATGCAACGCGAAGAACTTACCTACCTT 960  
 DB 928 GCACAAGCGGTGATATGTGATTAATTGATGCAACGCGAAGAACTTACCTACCTT 987  
 QY 961 GACATGTACGAATTTTCTAGAGATAGATTAGTCT--TGGGAACGCTAACACAGGTG 1017  
 DB 988 GACATGTGTGGAATCTGCTGAGAGGTGGAGTGTCTGAAAGAAACCGGCGCACAGGTG 1047  
 QY 1018 CTGCATGCGTGTGCTGAGCTCGTGTGAGATGTTGGTTAAAGTCCCGCAACGACGCA 1077  
 DB 1048 CTGCATGCGTGTGCTGAGCTCGTGTGAGATGTTGGTTAAAGTCCCGCAACGACGCA 1107  
 QY 1078 ACCCTGTGATTAATTGCCATCAATTTGTGGGCACTTTAATGACCTGCGGTGACAA 1137  
 DB 1108 ACCCTGTGCTTATGTTGCTAC-----GCAAGACACTCTAAGGAGACTGCGGTGACAA 1162  
 QY 1138 CCGAGGAAGGTGGGATGACCTCAAGTCTCTATGCGCTTAAAGGTAGGGCTTCAACG 1197  
 DB 1163 CCGAGGAAGGTGGGATGACCTCAAGTCTCTATGCGCTTAAAGGTAGGGCTTCAACG 1222  
 QY 1198 TAATCAATGCGCGTACAGAGGAGTGGCCAAACCGGAGGGGAGCTAATCTCAGAAAG 1257  
 DB 1223 TCATCAATGCGCGTACAGAGGAGTGGCCAAACCGGAGGGGAGCTAATCTCAGAAAG 1282  
 QY 1258 GCGTGTAGTCCGATCGAGTCTGCACTGCACTCGTGAAGTCGGAATCGTAGTAAT 1317  
 DB 1283 CATCTGATGTCGGATTGCACTTGCACTCGATGCACTGGAATCGTAGTAAT 1342  
 QY 1318 CCGGATCAGCATGTGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCGTCACAC 1377  
 DB 1343 CCGGATCAGCATGTGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCGTCACAC 1402  
 QY 1378 CATGGAAGGTGTTTACACGAGAGCGTAGTCTTAACCGTAAGAGGGGCTTGGCCACG 1437  
 DB 1403 CATGGAAGGTGTTTACACGAGAGCGTAGTCTTAACCGTAAGAGGGGCTTGGCCACG 1462  
 QY 1438 TGAATTCAATGACTGGGGTG 1457  
 DB 1463 TAGATTCAATGACTGGGGTG 1482

RESULT 29  
 ADB61689  
 ID ADB61689 standard; DNA; 1535 BP.  
 XX ADB61689;  
 AC ADB61689;  
 XX 04-DEC-2003 (first entry)  
 DT  
 XX  
 DE 16S rRNA of Burkholderia cepacia DNA sequence.  
 XX



Oy		1318	CGCGGATCAGCATGTCGGGTGTAATAACGTTCCCGGTCTTGTACAACGCCGCCGCACAC	1377
Dd		1343	CGCGGATCAGCATGTCGGGTGTAATAACGTTCCCGGTCTTGTACAACGCCGCCGCACAC	1402
Oy		1378	CATGGAGATGGGGTTTACCAGAAGCAGGTACTCAACCGTAAGAGAGCGCTTGCCACGG	1437
Dd		1403	CATGGAGATGGGGTTTACCAGAAGCGCTAAGCTAACCCGAAGAGACGGTCAACCAAG	1466
Oy		1438	TGAGATTTCATGACTGGGGTG	1457
Dd		1463	TAGATTTCATGACTGGGGTG	1482
RESULT 30				
ID	ADW12666		standard; DNA; 1530 BP.	
XX AC	ADW12666;			
XX DT	07-APR-2005	(first entry)		
XX DE	Variovorax paradoxus 16S rDNA, SEQ ID NO:1.			
XX KW	Stereoselective synthesis; enantiomeric enrichment;			
XX OS	beta-amino acid synthesis; cell culture; 16S ribosomal RNA; 16S rRNA; ds.			
XX PN	Variovorax paradoxus.			
XX XX	US2005009151-A1.			
PD	13-JAN-2005.			
PF	22-JUN-2004; 2004US-00875161.			
PR	10-JUL-2003; 2003US-0486032P.			
PR	02-SEP-2003; 2003US-0499622P.			
XX PA	(PHNA ) PHARMACIA CORP.			
XX PI	Chase M, Clayton R, Landis B, Banerjee A;			
DR	WPI; 2005-110890/12.			
XX PT	Stereoselective synthesis of beta-amino acid such as D-beta- or L-beta-			
PT	phenylalanine, by contacting amino donor and amino acceptor in presence			
XX PT	of beta-amino acid transaminase to form beta-amino acid enantiomer, from			
XX PT	amino acceptor.			
PS	Claim 74; SEQ ID NO 1; 44pp; English.			
XX CC	The invention relates to methods for the stereoselective synthesis and			
CC	for the enantioselective enrichment of a beta-amino acid or its salt. The			
CC	methods involve contacting an amino donor and an amino acceptor in the			
CC	presence of a stereoselective beta-amino acid transaminase to form a beta			
CC	-amino acid enantiomer (or its salt) from the amino acceptor. The			
CC	invention also relates to purified stereoselective D-beta-transaminases			
CC	derived from a microorganism of the genera Variovorax, Nocardi,			
CC	Comamonas, Rhodococcus or Pseudomonas; purified stereoselective L-beta-			
CC	transaminases derived from a microorganism of the genus Alcaligenes;			
CC	methods of selecting and culturing microorganisms which express a beta-			
CC	transaminase; purified cultures of Variovorax paradoxus or Rhodococcus			
CC	opacus; the 16S ribosomal RNA gene (rDNA) of Variovorax paradoxus or			
CC	Rhodococcus opacus (ADW12666 and ADW12667, respectively); and a method of			
CC	detecting these 16S rDNA sequences. The methods of the invention are			
CC	useful for stereoselectively synthesizing a beta-amino acid such as D-			
CC	beta-phenylalanine or L-beta-phenylalanine, or for enriching one or the			
CC	other enantiomer in a racemic mixture. In the examples of the invention,			
CC	microorganisms from soil samples from two different environments were			
CC	screened transaminase activity on DL-beta-phenylalanine. Ribosomal DNA			
CC	sequencing was used to facilitate the taxonomic identification of two			
CC	purified microorganisms from the environmental samples. The present			
CC	sequence represents the specifically claimed 16S rDNA sequence which was			

CC	Identified as originating from	Varilovorex paradoxus.
XX		
Sequence	1530 BP; 389 A; 354 C; 475 G; 310 T; 0 U; 2 Other;	
Query Match	78.1%; Score 1137.6; DB 14; Length 1530;	
Best Local Similarity	87.5%; Pred. No. 0;	
Matches 1278; Conservative	1; Mismatches 175; Indels 6; Gaps 3;	
QY	1	ATTGAACGCTGGCGGCATGCTTTTACACATGCAATCGAACCAGACACGATGCTTGCAT 60
DB	24	ATTGAACGCTGGCGGCATGCTTTTACACATGCAATCGAACCAGACACGATGCTTGCAT-- 81
QY	61	CTGGTGGCGAATGGCCGACCGGTGAGTAATGCAATCGAAGCTATCCAAAGAGGGGGATA 120
DB	82	CTGGCGCGAATGGCCGACCGGTGAGTAATGCAATCGAAGCTATCCAAAGAGGGGGATA 141
QY	121	ACGCATCGAAGATGTGTCTAATACCGCATATACTCTAAGAGAGAAAGAGGGATCGAAA 180
DB	142	ACGCAGCGAAGATGTGTCTAATACCGCATATACTCTAAGAGAGAAAGAGGGATCGAAA 201
QY	181	GACCTTGGCTTTTGGAGCGGCCGATGCTGATTAAGCTAGTGGTGGGTAAAGGCTTAC 240
DB	202	GACCTTGGCGAATGAGCGGCCGATGATTAAGTATGTTGGTGGGTAAAGGCTTAC 261
QY	241	CAAGCGCAACATCAATGATTTGTTGAGAGGACACACACGACCTGGGATCTGAGACAG 300
DB	262	CAAGCTTGCATCTGTAGCTGTCTGAGAGGACACACACGACCTGGGATCTGAGACAG 321
QY	301	GCCCAAGCTTCCATCGAGGAGGACAGAGTGGGGAATTTTGGACAAATGGCGCAAGCTGATC 360
DB	322	GCCCAAGCTTCCATCGAGGAGGACAGAGTGGGGAATTTTGGACAAATGGCGCAAGCTGATC 381
QY	361	CAGCAATGCGCGGTGAGTGAAGAAAGGCTTGGGTTGAAAGCTCTTCACTCGAAGAA 420
DB	382	CAGCATGCGCGGTGAGTGAAGAAAGGCTTGGGTTGAAAGCTCTTCACTCGAAGAA 441
QY	421	AAAGTTACGGTAATTAATTCGTGACTCATGACGATTCGACAGAAAGAACCGGCTTAC 480
DB	442	AAAGCTCTTTTCTAATAAAGGGCTAATGACGGTACGTAAGATTAAGAACCGGCTTAC 501
QY	481	TACGTGCGAGACGCGCGGTAAATACGTAGGGTGCAGAGCTTAATTCGAAATTAATCGAGCGT 540
DB	502	TACGTGCGAGACGCGCGGTAAATACGTAGGGTGCAGAGCTTAATTCGAAATTAATCGAGCGT 561
QY	541	AAAGGTCGCGAGCGCGCTTTGTAAGTCAGATGTGAATTCGCCGGCTTAACTGGGAAT 600
DB	562	AAAGGTGCGAGCGCGGTATGTAAGACATGTGTAATTCGCCGGCTTAACTGGGAATC 621
QY	601	TGCGTTTGAATCAACAGGCTTAAGATGTGGCAGAGGAGGTGGAAATTCATGTTAGAG 660
DB	622	TGCACTGTGACTGATAGCTAGAGTAAGGTAGAGGGGGAATGGAATTCGCGTGTAGAG 681
QY	661	TGAATGGTGTGAGATGATGGAAGAACATCGATGGGAGGACGCTCTGGGTTAACT 720
DB	682	TGAATGGTGTGATGATGGAAGAACATCGATGGGAGGACGCTCTGGGTTAACT 741
QY	721	GACGCTCATGCAAGAAAGCTGGGAGCAACAGGATTAGATCCCTGTAGTCCAGCC 780
DB	742	GACGCTCATGCAAGAAAGCTGGGAGCAACAGGATTAGATCCCTGTAGTCCAGCC 801
QY	781	CTAAACGATGTCACTAGTTGTTGGCCCTTATAGGCTTGGTAAACGATTAACCGGTA 840
DB	802	CTAAACGATGTCACTAGTTGTTGGGAATTCATCTTTCAGTAAAGAAAGCTTAACCGGTA 861
QY	841	AATTGACCGCTGGGAGAGTACGATGGCAAGATTAAATCTCAAGAAATTAAGCGGGAGCC 900
DB	862	AATTGACCGCTGGGAGAGTACGATGGCAAGATTAAATCTCAAGAAATTAAGCGGGAGCC 921
QY	901	GCACAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAATACTTAACCTTACCTT 960
DB	922	GCACAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGAATACTTAACCTTACCTT 981
QY	961	GACATGTACGGAATTTTCTAGAGATGATTAAGTACTCT--TCGGGAACCTTAACAGGTG 1017

DB 982 GACATGTACGGAATTCGCCAGATGGCTTAAGTCTCGAAGAAAGAACCTTAACACAGGTG 1041  
QY 1018 CTGCATGCGTGTCTGTCAGCTGTGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCA 1077  
DB 1042 CTGCATGCGTGTCTGTCAGCTGTGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCA 1101  
QY 1078 ACCCTGTCAATTAATGTCATCTTTGGTGGCACTTTAAAGACATCCCGGTGCAAA 1137  
DB 1102 ACCCTGTCAATTAATGTCATCTTTGGTGGCACTTTAAAGACATCCCGGTGCAAA 1160  
QY 1138 CCGGAGGAAGTGGGGATGACGTCAGTCTCTCATGAGCCCTTAAGGTAGGCTTCAACAG 1197  
DB 1161 CCGGAGGAAGTGGGGATGACGTCAGTCTCTCATGAGCCCTTAAGGTAGGCTTCAACAG 1220  
QY 1198 TAATACATAGCGCGCTACAGAGGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAAC 1257  
DB 1221 TCATACATAGCGCGCTACAGAGGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAAC 1280  
QY 1258 GCGTGTAGTCCGGATCGGAGTCTGCACTCGACTCCGTAAGTCCGATCTAGTAAAT 1317  
DB 1281 CAGTGTAGTCCGGATCGGAGTCTGCACTCGACTCCGTAAGTCCGATCTAGTAAAT 1340  
QY 1318 GCGGATCGAGCATGTGCGGGTGAATACGTTCCCGGGTCTTGAACACACCGCCGTACAC 1377  
DB 1341 CGTGATCGAATGTACAGGTGATACGTTCCCGGGTCTTGAACACACCGCCGTACAC 1400  
QY 1378 CATGGAGTGGGTTTCACAGAGAGCAGTATCTTAACCGTAAAGAGGGCGCTTGCCACG 1437  
DB 1401 CATGGAGGCGGTTTCGACAGAGATGTTAGCTTAACCCGAAAGAGGGCGATTAACACG 1460  
QY 1438 TGAGATTCTGACTGGGGTG 1457  
DB 1461 CAGGTTCTGTGACTGGGGTG 1480

RESULT 31  
ADB61687  
ID ADB61687 standard; DNA; 1485 BP.  
XX  
AC ADB61687;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE 16S rRNA of Bordetella parapertussis DNA sequence.  
XX  
KM enriching mRNA; high quality bacterial mRNA; bacterial gene expression;  
XX  
KM poly-A tail; mRNA purification; oligo-dT capture;  
XX  
KM prokaryote mRNA purification; bridging oligonucleotide; targeting region;  
KM capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;  
KM eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;  
KM 28S eukaryotic rRNA bridging oligonucleotide; de.  
XX  
OS Bordetella parapertussis.  
XX  
FN WO2003054162-A2.  
XX  
PD 03-JUL-2003.  
XX  
PF 19-DEC-2002; 2002WO-US041014.  
XX  
PR 20-DEC-2001; 2001US-00029397.  
XX  
PA (AMBI-) AMBION INC.  
XX  
PI Murphy GL, Whitley JP;  
XX  
DR WPI; 2003-663255/62.  
XX  
PT bridging or isolating targeted nucleic acids e.g. rRNA, involves using a  
PT region complementary to a targeted nucleic acid, and a capture  
PT oligonucleotide.

XX  
PS Claim 4; Page 172; 208pp; English.  
CC This invention relates to a novel method for isolating, depleting or  
XX separating a targeted nucleic acid, such as rRNA, from a sample  
CC comprising targeted and non-targeted nucleic acids. It effects a way of  
CC enriching for non-targeted nucleic acids such as mRNAs. Isolating  
CC sufficient quantities of high quality bacterial mRNA is a demanding  
CC process which impedes analysis of bacterial gene expression in the  
CC presence of host cells. A small percentage of bacterial mRNAs may be poly  
CC -A tailed, but these are targeted for degradation and tend to be  
CC unstable. As a result, the commonly employed method for mRNA purification  
CC with eukaryotic cells, oligo-dT capture, is ineffective. The present  
CC invention provides an alternative, more suitable method for mRNA  
CC purification from prokaryotes. The method of the invention comprises the  
CC incubation of a sample with a bridging oligonucleotide (containing a  
CC targeting region) and subsequently incubating with a capture  
CC oligonucleotide allowing the isolation of the target from the sample. The  
CC method is useful for depleting or isolating targeted nucleic acid, for  
CC example rRNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S  
CC or 18S or eukaryotic 28S rRNA, from a sample. The rRNA sequence may  
CC comprise any one of 64 fully defined sequences as given in the  
CC specification. The present sequence is that of a DNA sequence which  
CC represents the sequence of 16S rRNA of Bordetella parapertussis related  
CC to the invention.  
XX  
SQ Sequence 1485 BP; 369 A; 350 C; 478 G; 288 T; 0 U; 0 Other;  
Query Match 78.1%; Score 1137.2; DB 10; Length 1485;  
Best Local Similarity 88.0%; Pred. No. 0;  
Matches 1287; Conservative 0; Mismatches 163; Indels 12; Gaps 4;  
QY 1 ATTTGAACCTGGCGGATGCTTTTACACATGCAAGTCCGAACGGCAGCAGGATCTTGAT 60  
DB 1 ATTTGAACCTGGCGGATGCTTTTACACATGCAAGTCCGAACGGCAGCAGG--GCTTGGC 58  
QY 61 CTGGTGGGAGTGGGCGAGCGGTGATGATCGAATCGAATCTATCCAGAAAGGGGGGTA 120  
DB 59 CTGGTGGGAGTGGGCGAGCGGTGATGATCGAATCTATCCAGAAAGGGGGGTA 118  
QY 121 ACGCATCGAAGATGTGCTAATACCGCATATCTTAAAGAGAAAGCAGGGGATCGAA 180  
DB 119 ACTAGCGAAGACGGTGTAAATACCGCATATCGCCCTACGGGGGAAAGCGGGGACTTTG 178  
QY 181 GACCTTGGCTTTTGGAGCGGCCCATGTCTGATTAAGCTAAGTTGGGGTAAAGGCTTAC 240  
DB 179 GACCTTGGCATATTGAGACGGCCGATATCGAATTAGTTGGGGTAAAGGCTTAC 238  
QY 241 CAAGCGACGATCTAGTATTGTTGTCGAGAGGACGACGACACTGGGACTGAGACAG 300  
DB 239 CAAGCGACGATCTAGTATTGTTGTCGAGAGGACGACGACACTGGGACTGAGACAG 298  
QY 301 GCCCAGACTCTCTACCGGAGGAGCAGAGTGGGGAATTTTGGACATGAGCGCAAGCTGATC 360  
DB 299 GCCCAGACTCTCTACCGGAGGAGCAGAGTGGGGAATTTTGGACATGAGCGCAAGCTGATC 358  
QY 361 CAGCATATCCCGCGTGAAGTGAAGAGGCTTGGGTTGAAGCTCTTCACTCGAAGAGA 420  
DB 359 CAGCATATCCCGCGTGAAGTGAAGAGGCTTGGGTTGAAGCTCTTGGCAGGAAAGA 418  
QY 421 AAAGGTTACGGTAAATATCTGATCTCATGACGGATGAGCAGAAAGACCGGCTAAC 480  
DB 419 AAAGGTTACGGTAAATATCTGATCTCATGACGGATCTCGAAGATTAAGCACCGGCTAAC 478  
QY 481 TACGTGCGAGAGCGCGCGTAAATAGTGAAGGTGCAAGCGTTAATCGAATTAATCGGGCGT 540  
DB 479 TACGTGCGAGAGCGCGCGTAAATAGTGAAGGTGCAAGCGTTAATCGAATTAATCGGGCGT 538  
QY 541 AAAGGTTGCGAGCGCGCTTTGTAAGTCAATGTGAATCCCGGGCTTAACTCGGGAAT 600  
DB 539 AAAGGTTGCGAGCGCGCTTTGTAAGTCAATGTGAATCCCGGGCTTAACTCGGGAAT 598  
QY 601 TGGCTTGAATCTCAAGAGCTTAGAGTGTGGCAGGAGGTGAATTCATGTGTAAG 660

Db	599	TGCAATTTTAACTACCGGGCTAGAGTGTCTAGAGGAGAGTGGAAATTCGCGCTGTAGCAAG	658
Qy	661	TGAAATGCGTAGAGATATATGAAAGAACATGATGAGTGGCCAAAGGCAAGCTCTTGGGTTAAACAT	720
Db	659	TGAAATGCGTAGATATATGCGAGAGAAACACGATGGCGAAGGCAAGCTCTTGGGTTAAACAT	718
Qy	721	GACGCTCATGCAAGAAAGCGTGGGGAGCAACAGATTAGATACCTGGTAGTCCAGCC	780
Db	719	GACGCTCATGCAAGAAAGCGTGGGGAGCAACAGATTAGATACCTGGTAGTCCAGCC	778
Qy	781	CTAAACGATGTCACCTAGTTGTTGGGCTTATTAAGCTTGGTAAACGAGCTAACGGGTGA	840
Db	779	CTAAACGATGTCACCTAGTTGTTGGGCTTATTAAGCTTGGTAAACGAGCTAACGGGTGA	838
Qy	841	AGTTAGACGCGCTGGGGAGTAGCGGTCCGACAATTAACCAAGGAATTGACCGGGACCC	900
Db	839	AGTTAGACGCGCTGGGGAGTAGCGGTCCGACAATTAACCAAGGAATTGACCGGGACCC	898
Qy	901	GCACAAGCGGTGATTTATGTGGAATTAATTGCATGACACGCGAAAACTTTACTTACCTT	960
Db	899	GCACAAGCGGTGATTTATGTGGAATTAATTGCATGACACGCGAAAACTTTACTTACCTT	958
Qy	961	GACATGTACCGAATTTTCTAGAGATAGATTAGTGTCT---TCGGGAACGCTTAACACAGGTG	1010
Db	959	GACATGTCTGGAATATCCCGAAGAGATTGGGAGTCTCGCAAGGAACCCGGAACACAGGTG	1010
Qy	1018	CTGCATGCGCTGATCTGACGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA	1070
Db	1019	CTGCATGCGCTGATCTGACGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA	1070
Qy	1078	ACCTCTTGCTAATTAATTGCAATCATCTTTGTGGCACTTTAATGAGACTGCGGTGACAA	1133
Db	1079	ACCTCTTGCTAATTAATTGCTAATGCAAA-----GGGCACTCTAATGAGACTGCGGTGACAA	1133
Qy	1138	CCGAGAGAAAGTGGGGATGACGTCAAGTCTCATGCGCTTTAATGGGTAGGGCTTACACAG	1199
Db	1134	CCGAGAGAAAGTGGGGATGACGTCAAGTCTCATGCGCTTTAATGGGTAGGGCTTACACAG	1199
Qy	1198	TAAATCAATGGGCGGTGACAGAGGGTTGGCCAAACCGGAGGGGGAGCTAATCTCAGAAAGC	1255
Db	1194	TAAATCAATGGTGTGGGACAGAGGGTGGCCAAACCGGAGGGGGAGCCAAATCCAGAAAGC	1255
Qy	1258	GCGTGTAGTCCGGATCGAGTCTGCAACTGCACTCCGTGAAGTCCGAATCGCTAGTAAT	1311
Db	1254	GCGTGTAGTCCGGATCGAGTCTGCAACTGCACTCCGTGAAGTCCGAATCGCTAGTAAT	1311
Qy	1318	CGCGGATCAGCATGTGCGGTGAATACGTTCCCGGGCTTTGTACACACCGCCGTCAAC	1373
Db	1314	CGCGGATCAGCATGTGCGGTGAATACGTTCCCGGGCTTTGTACACACCGCCGTCAAC	1373
Qy	1378	CATGGAGTGGGTTTACACAGAGCAGTGTCTAACCCGTA--GGAGGGCGCTTGGCAC	1434
Db	1374	CATGGAGTGGGTTTACACAGAGTGTGTAACTTAACCGCACAGGAGGGGGGCGATTACAC	1434
Qy	1436	GGTGAGATTCAATGACTGGGGTG 1457	
Db	1434	GGTAGATTCAATGACTGGGGTG 1455	
RESULT 32			
ABZ69299			
ID	ABZ69299	standard; DNA; 1400 BP.	
AC	ABZ69299;		
XX	11-AUG-2003	(first entry)	
XX	J 11vidum 16S ribosomal RNA gene fragment #3.		
DE	Osteoarthritis; antibacterial; Jantchinobacterium; 16S RNA; gene;		
XX	osteopathic; antiarthritic; analgesic; anti-inflammatory; ds.		
XX			

Query Match	Best Local Similarity	Matches 1257; Conservative	Score 77.8%	Pred. No. 0; Mismatches 140; Indels 10; Gaps 3	DB 8; Length 1400;
XX OS	Janthinobacterium lividum.				
XX FN	WO2002102384-A1.				
XX PD	27-DEC-2002.				
XX PF	17-JUN-2002; 2002WO-GB002771.				
XX PR	15-JUN-2001; 2001GB-00014672.				
XX PA	(ORTH-) ORTHOGENICS AS.				
XX PA	(GARD/) GARDNER R.				
XX PI	El-Gewely MR;				
XX DR	WPI; 2003-17519/17.				
XX PT	Use of an antibacterial agent in the manufacture of a medicament for treating osteoarthritis.				
XX PS	Example 1; Page 49-51; 89pp; English.				
XX CC	The present invention relates to the use of an antibacterial agent in the production of a treatment for osteoarthritis. The bacteria causing osteoarthritis is Janthinobacterium lividum. The present sequence is a fragment of the 16S RNA coding sequence from J. lividum shown in the exemplification of the invention				
XX CC	Sequence 1400 BP; 371 A; 309 C; 432 G; 288 T; 0 U; 0 Other;				
XX SQ					
QY	27 CATGCAAGTCGAAACGGACGACAGATGCTTGATCTGTGTGCGGACGAGCGAGCGGTGAG	86			
DB	1 CATGCAAGTCGAAACGGACGACAGG--GCTTCTCTGTGTGCGGAGTGGCGGACGGGTGAG	58			
QY	87 TAATGATCTCGAAACGTATTCAGAAAGAGGGGGTAAACGATCGAAGAGTGTCTAATACCG	146			
DB	59 TAATATATCGGAACGTACCTTAGATGGGGGGATTAAGTAGCGAAAGTTACGCTAATATACCG	118			
QY	147 CATATATCTTAAGAGGAAAGCAGGGGATTCGAAACCTTGCGCTTTTGGACGGCGCAT	206			
DB	119 CATACGATCTTAAGATGAAAGTGGGGGATTCGCAAGCCTCATGCTGTGTGACCGGCGCAT	178			
QY	207 GCTCATTTAGCATGATTTGGTGGGGGTAAAGGCTTACCAAGGCGACGATCACTAGTTGGTCTG	266			
DB	179 ATCTATATTGATGATTTGGTGGGTAAAGCTTACCAAGGCGATCACTAGTTGGTCTG	238			
QY	267 AGAGACGACGACGACCACTGGGAGCTGAGACACCGGCCAGACTCTCTACCGGAGCGACAG	326			
DB	239 AGAGACGACGACGACCACTGGAATCTGAGACACCGTCCAGACTCTCTACCGGAGCGACAG	298			
QY	327 TGGGGAATTTTGGACATGAGGCGCAAGGCTGATCCAGCAATGCGCGTGAAGTGAAGAAG	386			
DB	299 TGGGGAATTTTGGACATGAGGCGCAAGGCTGATCCAGCAATGCGCGTGAAGTGAAGAAG	358			
QY	387 CCTTGGGGTGTGTAAGGCTTTTCACTGCGAAGAAAGGTTACGTAATTAATTCGAGACT	446			
DB	359 CCTTGGGGTGTGTAAGGCTTTTGTCAAGGAAAGAAACGGTGAAGCTTAATTAATTCCTTGGCT	418			
QY	447 CATGACGGTATGACGAAAGAAAGACCGGCTTAATCACTGTCGACGACCGCGGTAAATACG	506			
DB	419 AATGACGGTATCTGAAAGAAATAGCACCGGCTTAATCACTGTCGACGACCGCGGTAAATACG	478			
QY	507 TAGGGTGCAGAGCTTAATCGAATTACTGGGCGTAAAGGGTGGCGAGGGCGCTTTGTAA	566			
DB	479 TAGGGTGCAGAGCTTAATCGAATTACTGGGCGTAAAGGGTGGCGAGGGCGCTTTGTAA	538			
QY	567 TCAGATGTGAATTCCTCGGGGCTTAACCTGGGAATTGCGTTGAAATCAACGGCTAGAGT	626			
DB	539 TCGATGTGAATTCCTCGGGGCTTAACCTGGGAATTGCAATTCAGACTTCGAAGGCTAGAT	598			



```
QY 627 GTGGGAGAGGAGTGAATTCATGTTAGCAGTGAATGGGTAGAGATATGAAGAC 686
XX |||||
OS 599 CTGGCAGAGGGGGTGAATTCAGTGAAGTGAATGGGTAGAGATATGAAGAC 658
XX |||||
PN 687 ATCGATGGCGAAGCAGCTCTGGGTAACTGACGCTCATGACGAAACGTTGGGA 746
XX |||||
PD 659 ACCGATGGCGAAGCAGCTCTGGGTCAAGATTGACGCTCATGACGAAACGTTGGGA 718
XX |||||
PF 747 GCAAAAGAGATTAGATACCTGTTAGTCAACGCTTAAACGATGTCATTAAGTTGGG 806
XX |||||
PR 719 GCAAAAGAGATTAGATACCTGTTAGTCAACGCTTAAACGATGTCATTAAGTTGGG 778
XX |||||
PA 807 CCTTTATTGGCTTGGTGAAGAGCTAACGCGTGAAGTTGACCGCTGGGGAGTACGCTG 866
XX |||||
PI 779 TTTTATTGACTTGGTGAAGAGCTAACGCGTGAAGTTGACCGCTGGGGAGTACGCTG 838
XX |||||
DR 867 CAAGATTAACTCAAGAGATTGACGCGGAGCCGCAAGCGGTTGATATGTTGATTA 926
XX |||||
PT 839 CAAGATTAACTCAAGAGATTGACGCGGAGCCGCAAGCGGTTGATATGTTGATTA 988
XX |||||
PT 927 ATTCATGCAACGCGAAAACTTACCTTACCTTGAATGACGAAATTTCTAGAGATA 986
XX |||||
PS 899 ATTCATGCAACGCGAAAACTTACCTTACCTTGAATGACGAAATTTCTAGAGATA 958
XX |||||
XX 987 GATTATGCT---TCGGGAACGCTAACAGAGGTGTCATGCTGTCAGCTGCTGTC 1043
XX |||||
CC 959 GAGGATGCTCGAAGAGAACGATACAGAGTGTGTCATGCTGTCAGCTGCTGTC 1018
XX |||||
CC 1044 GTGAGATGTTGGGTTAAGTCCGCAACGAGCGCAACCTTGTCAATTAATTCATCATTT 1103
XX |||||
CC 1019 GTGAGATGTTGGGTTAAGTCCGCAACGAGCGCAACCTTGTCAATTAATTCATCATTT 1073
XX |||||
CC 1104 GGTGGGCACTTTAATGAGACTGCGGTGACAAACCGAGGAAGTGGGGATGACGTCAA 1163
XX |||||
CC 1074 GAAAGGGCACTTTAATGAGACTGCGGTGACAAACCGAGGAAGTGGGGATGACGTCAA 1133
XX |||||
CC 1164 GTTCCTCAGGCGCTTATGAGGTGGGCTTACACGTAATACANTGGCGGCTACAGAGGTT 1223
XX |||||
CC 1134 GTTCCTCAGGCGCTTATGAGGTGGGCTTACACGTAATACANTGGCGGCTACAGAGGTT 1193
XX |||||
CC 1224 GCCAACCCGCGAGGGGAGCTATCTCAGAAAGCGCGTGTGTCGCGATCGGAGTGTGC 1283
XX |||||
CC 1134 GCCAACCCGCGAGGGGAGCTATCTCAGAAAGCTGATGTCGCGATTTGATGTCGC 1253
XX |||||
CC 1284 AACTGACTCGGTGAAGTGGATGCTGTAATGCGCGATCAAGATGTCGCGTGAATA 1343
XX |||||
CC 1254 AACTGACTCGGTGAAGTGGATGCTGTAATGCGCGATCAAGATGTCGCGTGAATA 1313
XX |||||
CC 1344 CGTCCCGGGGCTTGTACACACGCGCCGTCACACACATGGGAGTGGGTTTACACAGAGCA 1403
XX |||||
CC 1314 CGTCCCGGGGCTTGTACACACGCGCCGTCACACACATGGGAGCGGGTTTACAGAGATTA 1373
XX |||||
CC 1404 GGTAGTCTAACCGTAAGAGAGCGCGCTT 1430
XX |||||
CC 1374 GGTAGTCTAACCGTAAGAGAGCGCGCTT 1400
XX |||||

RESULT 33
ADB61688
ID ADB61688 standard; DNA; 1464 BP.
XX
AC ADB61688;
XX
DT 04-DEC-2003 (first entry)
XX
DE 16S rRNA of Bordetella pertussis DNA sequence.
XX
KM enriching mRNA; high quality bacterial mRNA; bacterial gene expression;
KM poly-A tail; mRNA purification; oligo-dT capture;
KM prokaryote mRNA purification; bridging oligonucleotide; targeting region;
KM capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;
KM eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;
```

```
KM 28S eukaryotic rRNA bridging oligonucleotide; ds.
XX
OS Bordetella pertussis.
XX
PN WO2003054162-A2.
XX
PD 03-JUL-2003.
XX
PF 19-DEC-2002; 2002WO-US041014.
XX
PR 20-DEC-2001; 2001US-00029397.
XX
PA (AMBI-) AMBION INC.
XX
PI Murphy GL, Whitley JP;
XX
DR WPI; 2003-663255/62.
XX
PT Depleting or isolating targeted nucleic acids e.g. rRNA, involves using a
PT bridging oligonucleotide comprising bridging region and a targeting
PT region complementary to a targeted nucleic acid, and a capture
PT oligonucleotide.
XX
PS Claim 4; Page 173; 208pp; English.
XX
CC This invention relates to a novel method for isolating, depleting or
CC separating a targeted nucleic acid, such as rRNA, from a sample
CC comprising targeted and non-targeted nucleic acids. It effects a way of
CC enriching for non-targeted nucleic acids such as mRNAs. Isolating
CC sufficient quantities of high quality bacterial mRNA is a demanding
CC process which impedes analysis of bacterial gene expression in the
CC presence of host cells. A small percentage of bacterial mRNAs may be poly
CC -A tailed, but these are targeted for degradation and tend to be
CC unstable. As a result, the commonly employed method for mRNA purification
CC with eukaryotic cells, oligo-dT capture, is ineffective. The present
CC invention provides an alternative, more suitable method for mRNA
CC purification from prokaryotes. The method of the invention comprises the
CC incubation of a sample with a bridging oligonucleotide (containing a
CC targeting region) and subsequently incubating with a capture
CC oligonucleotide allowing the isolation of the target from the sample. The
CC method is useful for depleting or isolating targeted nucleic acid, for
CC example rRNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S
CC or 18S or eukaryotic 28S rRNA, from a sample. The rRNA sequence may
CC comprise any one of 64 fully defined sequences as given in the
CC specification. The present sequence is that of a DNA sequence which
CC represents the sequence of 16S rRNA of Bordetella pertussis related to
CC the invention.
XX
SQ Sequence 1464 BP; 361 A; 344 C; 468 G; 285 T; 0 U; 6 Other;
XX

Query Match 77.7%; Score 1131.4; DB 10; Length 1464;
Best Local Similarity 87.8%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 167; Indels 10; Gaps 3;

QY 1 ATTTGACGCTGGCGGATGCTTTACATGACGAGTCGACCGGACGAGTCTTGAT 60
XX |||||
DB 27 ATTGAACCTGCGCGGAGTCTTTACATGACGAGTCGACCGGACGAGTCTTGAT 84
XX |||||

QY 61 CTGTGGCGAGTGGCGGAGCGGGTGAATCATCGGACGATATCCAGAGAGGGGGTGA 120
XX |||||
DB 85 CTNGTGGCGAGTGGCGGAGCGGGTGAATCATCGGACGATATCCAGAGAGGGGGTGA 144
XX |||||

QY 121 ACGCATGGAAGAGTGTGTATACCGCATATATCTTAAGAGGAGGAGGAGGAGTCA 180
XX |||||
DB 145 ACTACGCGAAGCGTGAATATACCGCATATATCTTAAGAGGAGGAGGAGGAGTCA 204
XX |||||

QY 181 GACCTTGCGCTTTTGAAGCGCGCGATGCTGATTAAGTTGGTGGGTTAAAGCCTTAC 240
XX |||||
DB 205 GGCCTCGCATTAATGAGAGCGCGCGATATCGAATTAAGTGTGGTGAAGCGGCTTAC 264
XX |||||

QY 241 CAAGCGCAGATCATGATGTTGTTGAGAGGACGACCAACGACACTGGAGTGAACACG 300
XX |||||
DB 265 CAAGCGCAGATCATGATGTTGTTGAGAGGACGACCAACGACACTGGAGTGAACACG 324
XX |||||
```



```

OY 301 GCCCAGACTCTTACCGGAGGAGCAGATGGGGAATTTTGGACAATGGGCGCAAGCTGATC 360
DB 325 GCCCAGNCTCTTACCGGAGGAGCAGATGGGGAATTTTGGACAATGGGCGCAAGCTGATC 384
OY 361 CAGCAATGCGGGGTGAGTGAAGAAGGCTTTGGGTTGTAAGCTCTTTCAATCGAANA 420
DB 385 CAGCCATCCCGGTGCGATGAGGAGGCTTGGGTTGTAAGCACTTTGGCAGGAAGA 444
OY 421 AAAGGTACGTAATATCTGTAATCGTATCGATCGGATCGACAGAAAGACCGGCTAAC 480
DB 445 AACGGACGGGCTAATATCTGTCGACTGACGGTACTCTGACAAATACGACCGGCTAAC 504
OY 481 TACGTCCAGCAGCCCGGTAATACGTAGGGTGCAAGCGTTAATCGAAATTACTGGCGCT 540
DB 505 TACGTCCAGCAGCCCGGTAATACGTAGGGTGCAAGCGTTAATCGAAATTACTGGCGCT 564
OY 541 AAAGGTGCGCAGCGGCTTTTGAATGATGATGTAATCCCGGGCTTAACCTGGGAAT 600
DB 565 AAAGCGTGGCAGCGGCTTCGAAAGAAAGATGTGAATCCCGGGCTTAACCTGGGAAT 624
OY 601 TCGCTTGAACATCAAGGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTAGCAG 660
DB 625 TCGATTTTAACTACCGGGCTAGAGTGTGTCAAGGAGGTGAATTCGCGCTGTAGCAG 684
OY 661 TGAATGCGTAGAGATATGAGAAAGACATCGATGGCGAAGCGACCTCTGGGTTAACT 720
DB 685 TGAATGCGTAGAGATATGAGAAAGACATCGATGGCGAAGCGACCTCTGGGTTAACT 744
OY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTGTGTAGTCCAGCC 780
DB 745 GACGCTCATGACGAAAGCGTGGGAGCAAAAGATTAGATACCTGTGTAGTCCAGCC 804
OY 781 CTAAAGATGTCAATAGTTGTGGGCTTATTAAGCTGTGTGTACGAAGCTAACGGGTGA 840
DB 805 CTAAAGATGTCAATAGCTGTGTGGGCTTTCGGGCTTGTGTAGCGAGTTAAAGCTGA 864
OY 841 AGTTGACCGCTGGGAGTAGTGGTCCCAAGATTAAACTCAAGAAATTGACCGGAGCC 900
DB 865 AGTTGACCGCTGGGAGTAGTGGTCCCAAGATTAAACTCAAGAAATTGACCGGAGCC 924
OY 901 GCAAGAGCGGTGATATGTGATTTAATTGATGCAACGCGAAAAACCTTACCTCTT 960
DB 925 GCAAGAGCGGTGATATGTGATTTAATTGATGCAACGCGAAAAACCTTACCTCTT 984
OY 961 GATATGTAGGAATTTTCTAGATATAGTAACT---TCGGAAAGCGTAACACAGGTG 1017
DB 985 GATATGTGTGAATCTCCGAAGAGATTTGGAGTGTCTCGCAAGAGAACCGGAACACAGGTG 1044
OY 1018 CTGCATGAGCTGTGTCAGCTCGTGTGCTGATGATTTAGGTTAAGTCCCGCAAGAGCGCA 1077
DB 1045 CTGCATGAGCTGTGTCAGCTCGTGTGCTGATGATTTAGGTTAAGTCCCGCAAGAGCGCA 1104
OY 1078 ACCCTGTCTAATTAATTGCAATCTTTGGTGGCACTTTAATAGACTGCGCGGTGACAA 1137
DB 1105 ACCCTGTCTAATTAATTGCAATCTTTGGTGGCACTTTAATAGACTGCGCGGTGACAA 1159
OY 1138 CCGGAGGAAGGTGGGAGTAGACGTCAACTCTCATGCGGCTTAAAGGAGGCTTCAACAG 1197
DB 1160 CCGGAGGAAGGTGGGAGTAGACGTCAACTCTCATGCGGCTTAAAGGAGGCTTCAACAG 1219
OY 1198 TAATACAAATGGCGGTACAGAGGGTGTCCAAACCGCAGAGGGGAGTAATCTCAGAAAGC 1257
DB 1220 TCATACAAATGGTGGGAGCAGAGGGTGTGCAACCGCAGAGGGGAGGCCAATCTCCAGAAAC 1279
OY 1258 GCGTGTAGTTCGGGATCGAGTCTGCAACTGCACTCGTGAAGTCCGAATCGCTAGTAAT 1317
DB 1280 CCGGTGTAGTTCGGGATCGAGTCTGCAACTGCACTCGTGAAGTCCGAATCGCTAGTAAT 1339
OY 1318 CCGGATCAGCATGTCCCGGTGAATTAAGTTCGCGGCTTGTGTACACACCGCGCTCACAC 1377
DB 1340 CCGGATCAGCATGTCCCGGTGAATTAAGTTCGCGGCTTGTGTACACACCGCGCTCACAC 1399

```

```

OY 1378 CATGGAGTGGGTTTACACAGAGCAGTAGTCTAACCGTAAGAGGGCGCTTCCACGG 1437
DB 1400 CATGGAGTGGGTTTACACAGAGCAGTAGTCTAACCGTAAGAGGGCGCTTACACGG 1459
OY 1438 TGAGA 1442
DB 1460 TAGGA 1464

RESULT 34
AAC86023
ID AAC86023 strand; cDNA; 1478 BP.
XX
XX AAC86023;
AC
XX 29-AUG-2001 (first entry)
DT
XX
DE R. purpureus 16S rDNA.
XX
XX 16S rDNA; polyphosphate accumulating organism; PAO; probe; primer;
KW detection; phosphorus; waste water; sludge; ss.
XX Rhodocyclus purpureus.
OS
XX MO20014659-A1.
PN
XX 28-JUN-2001.
PD
XX
PF 28-DEC-2000; 2000WO-AU001611.
PR 23-DEC-1999; 99AU-00004867.
XX
XX (CRCM-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.
PA
XX Hugenholz P, Crocetti GR, Tyson GW, Blackall LJ;
PI WPI; 2001-408656/43.
XX
XX Novel oligonucleotide probe or primer useful for detecting polyphosphate
PT accumulating organism in a sample, comprises a sequence that is unique to
PT 16S rDNA of polyphosphate accumulating organisms.
XX
XX Claim 4; Fig 3; 54pp; English.
PS
XX
XX The sequences given in AAC86021-30 represent 16S rDNA sequences from
XX polyphosphate accumulating organisms (PAOs). Sequences which are unique
XX to these 16S rDNA sequences are used to create a probe or primer for
XX detecting the relevant organisms. The primer/probe sequences are useful
XX for detecting PAO cells in a sample, by treating cells in the sample to
XX fix cellular contents, contacting fixed cells with the primer/probe which
XX is labelled with a radiolabel, a reporter group or a hapten, under
XX conditions which allow the probe to hybridize with 16S rRNA within the
XX fixed cell, removing unhybridized probe from the fixed cells, and
XX detecting the labeled probe-RNA hybrid by fluorescence in situ
XX hybridization. The primer/probe sequences are useful for identifying PAOs
XX that are capable of biologically removing phosphorus from waste water.
XX Rapid assessment of the presence of a number of PAOs in a waste water
XX sample, can be done using the primer/probe sequences. They allow quick
XX and convenient assessment of whether a sludge or waste water sample
XX includes PAOs and allows quantitation of PAO cells in samples
XX
XX Sequence 1478 BP; 367 A; 345 C; 470 G; 289 T; 0 U; 7 Other:
SQ
XX
XX Query Match 77.6%; Score 1130.6; DB 4; Length 1478;
XX Best Local Similarity 87.4%; Pred. No. 0;
XX Matches 1265; Conservative 0; Mismatches 176; Indels 7; Gaps 3;
OY 1 ATTGAACGCTGCGGCGGATGCTTTACATGCAAGTGAACGCGAGCAGCAGGATCTTGAT 60
DB 29 ATTGAACGCTGCGGCGGATGCTTTACATGCAAGTGAACGCGTA--ACGGAGCTTCGGG 86
OY 61 CTGCTGCCAGATGGCGGAGCGGTGAATGATGCAATCGGAACGTATCCGAAGAGGAGGATA 120

```

Db 87 GCCCGACGAGTGGCGAAGCGGTGAGTAATGCAATCGGAACATGCCCTGAAGTGGGGGATA 146  
 QY 121 ACGCATCGAAGATGTGTAATACCGCATATATCTTAAGAGAGAAAAGCAGGGAGTCGAAA 180  
 Db 147 ACGTAGCGAAGATACGCTAATATACGCAATATCTGTGAGCAGAGAAAAGCAGGGAGACTTCG 206  
 QY 181 GACCTTGGCCTTTTGGAGCGGCCGATGTCGATTAAGCTAGTGTGGTGGGTTAAAGGCTAC 240  
 Db 207 GGCCTTGGCCTTTTGGAGCGGCCGATGTCGATTAAGCTAGTGTGGTGGGTTAAAGGCTAC 266  
 QY 241 CAAGCGACGATCAAGTGTGCTGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300  
 Db 267 CAAGCGACGATCCGTAAGCGGGCTGAGAGATGATCCGCCACACTCGGAGCTGAGAGCAG 326  
 QY 301 GCCCAGACTCTTACCGGAGGAGCAGCAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360  
 Db 327 GCCCAGACTCTTACCGGAGGAGCAGCAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 386  
 QY 361 CAGCATGCGCGGTGAGTGAAGAGCGCTTCGGGTGTGAAGCTCTTTCAGTTCGAGAGAA 420  
 Db 387 CAGCCATGCGCGGTGAGTGAAGAGCGCTTCGGGTGTGAAGCTCTTTCAGTTCGAGAGAA 446  
 QY 421 AAGGTTACGTAATATCTGTAATCATGACGGTATCGACAGAAAGAACACCGGCTAAC 480  
 Db 447 AATCGGGTTTCTTAATACGGAACCCGGATGACGGTACCCGAGAGAAAGACACCGGCTAAC 506  
 QY 481 TACGTCCAGCAGCCCGGTAATACGTAGGGTTCAGAGCTTAATCGGAAATTAATCTGGGCGT 540  
 Db 507 TACGTCCAGCAGCCCGGTAATACGTAGGGTTCAGAGCTTAATCGGAAATTAATCTGGGCGT 566  
 QY 541 AAGAGGTGGCAGCGCGCTTTGTAAAGTCAAGTGTGAATCCCGGGCTTAACCTGGGAAT 600  
 Db 567 AAGAGGTGGCAGCGCGCTTTGTAAAGTCAAGTGTGAATCCCGGGCTTAACCTGGGAAT 626  
 QY 601 TCGGTTTGAACCTACAGAGCTAGAGTGTGAGCAGAGGAGGAGTGGAAATTCATGTGAGCAG 660  
 Db 627 TCGGTTTGTGACTGCAACACTAGTAGAGTACGCGAGAGGGGGGTGGAAATTCACGTTAGCAG 686  
 QY 661 TGAATGCGTAGATATGAAAGACATGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
 Db 687 TGAATGCGTAGATATGAGAGAGACATGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 746  
 QY 721 GACGCTCAGCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
 Db 747 GACGCTCAGCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 806  
 QY 781 CTAAACGATGTCAACTAGTGTG--GGCCTTAATTAAGCTTGGTAAAGAGCTAACGCGT 838  
 Db 807 CTAAACGATGTCAACTAGTGTG--GGCCTTAATTAAGCTTGGTAAAGAGCTAACGCGT 866  
 QY 839 GAAAGTTGACCGCTGGGAGTACGTCGCAAGATTAACTCAAGAGAAATGACGGGGAC 898  
 Db 867 GAAAGTTGACCGCTGGGAGTACGTCGCAAGATTAACTCAAGAGAAATGACGGGGAG 926  
 QY 899 CCGCACAACCGGTGATTAATGAGATTAATGAGTCAACGAGAAACCTTAACCTAACCC 958  
 Db 927 CCGCACAACCGGTGATTAATGAGATTAATGAGTCAACGAGAAACCTTAACCTAACCC 986  
 QY 959 TTGACATGAGGAAATTTTCTAGAGATAGATTAGTGC--TTGCGGAAAGCTTAACACAG 1015  
 Db 987 TTGACATGAGGAAATTTTCTAGAGATAGATTAGTGC--TTGCGGAAAGCTTAACACAG 1046  
 QY 1016 TGTCTGATGCTGTCTGAGCTGCTGAGTGTGAGATTTGGGTTAAGTCCCGCAACGAGG 1075  
 Db 1047 TGTCTGATGCTGTCTGAGCTGCTGAGTGTGAGATTTGGGTTAAGTCCCGCAACGAGG 1106  
 QY 1076 CAACCTTGTCAATTAATTCGATGATTTGGGAGCACTTAAATGAGTCCCGGAGACA 1135  
 Db 1107 CAACCTTGTCAATTAATTCGATGATTTGGGAGCACTTAAATGAGTCCCGGAGACA 1166  
 QY 1136 AACCGGAGAGGAGTGGAGTGAAGTCAAGTCTCATGAGCTTATGGTGAAGGCTTACACA 1195  
 Db 1167 AACCGGAGAGGAGTGGAGTGAAGTCAAGTCTCATGAGCTTATGGTGAAGGCTTACACA 1226

QY 1196 CGTAATACATGGCGCGCTTACAGAGAGGTTGCCAACCCCGAGAGGGAGGCTAATCTCAGAAA 1255  
 Db 1227 CGTCAATACATGGTTCGCTCAATAGGTTGCAAAACCCGAGAGGGAGGCTAATCCAGAAA 1286  
 QY 1256 GCGCGTCTGAGTCCGAGATCGAGAGTCTGCACTCGAAGTGGAAATCGCTAGTA 1315  
 Db 1287 GCGCATGCTAGTCCGAGATTGAGTGTGCACTCGAAGTGGAAATCGCTAGTA 1346  
 QY 1316 ATCGCGATTCAGCATGTGCGGTGAATATCTTCCGGGCTTTGTACACACCGCCCTCAC 1375  
 Db 1347 ATCGCGATTCAGCATGTGCGGTGAATATCTTCCGGGCTTTGTACACACCGCCCTCAC 1406  
 QY 1376 ACCATGGAGTGGGTTTACACAGAGAGTATGCTAACCGTAAGAGAGGAGGCTTGCAC 1435  
 Db 1407 ACCATGGAGTGGGTTTACACAGAGTATGCTAACCGTAAGAGAGGAGGCTTGCAC 1466  
 QY 1436 GGTGAGATTTCAT 1447  
 Db 1467 GGCACGTTTCGT 1478

RESULT 35

AB269300  
ID AB269300 standard; DNA; 1400 BP.

AC AB269300;

DT 11-AUG-2003 (first entry)

DE J lividum 16s ribosomal RNA gene fragment #4.

KM Osteoarthritis; antibacterial; Jantihnobacterium; 16s RNA; gene;

KW osteopathic; antiarthritic; analgesic; anti-inflammatory; de.

OS Jantihnobacterium lividum.

PN WO2002102384-A1.

PD 27-DEC-2002.

PF 17-JUN-2002; 2002WO-GB002771.

PR 15-JUN-2001; 2001GB-00014672.

PA (ORTH-) ORTHOSENICS AS.

PI (GARD/) GARDNER R.

PI El-Gewely MR;

DR WPI; 2003-175199/17.

PT Use of an antibacterial agent in the manufacture of a medicament for

PT treating osteoarthritis.

PS Example 1; Page 52-54; 89pp; English.

CC The present invention relates to the use of an antibacterial agent in the  
 CC production of a treatment for osteoarthritis. The bacteria causing  
 CC osteoarthritis is Jantihnobacterium lividum. The present sequence is a  
 CC fragment of the 16s RNA coding sequence from J. lividum shown in the  
 CC exemplification of the invention

XX Sequence 1400 BP; 370 A; 310 C; 433 G; 287 T; 0 U; 0 Other;

SQ Query Match 77.5%; Score 1129.8; DB 8; Length 1400;

Best Local Similarity 89.2%; Pred. No. 0; Mismatches 142; Indels 10; Gaps 3;

QY 27 CATGCAAGTCGAACGGCAGCAGCAGATGCTTGCATCTGGTGGGAGATGGCGGAGGCTGAG 86  
 Db 1 CATGCAAGTCGAACGGCAGCAGCAGG--GCTTGTCTGTGGTGGGAGATGGCGGAGGCTGAG 58

QY 87 TAAATGATCGAAGCGTATCCAGAAAGGAGGAGTAAACGATCGAAGATGTCTAATACCG 146  
DB 59 TAAATATATCGAAGCGTATACCTTAGAGTGGGAGATTAAGTAGCGAAGATTACGCTAATACCG 118  
QY 147 CATATATCTTAAGAGAGAAAGACGAGGAGATCGAAGAGCCTTGCGCTTTTGGAGCGGCGCAT 206  
DB 119 CATACATCTTAAGAGATGAAGAGTGGGAGATCGAAGAGCCTTCACTGTGTGAGCGGCGCAT 178  
QY 207 GCTGATTAAGCTAGTGTGGGTAAAGGCTTCAAGGCGAGCATAGTATGTTGCTCTG 266  
DB 179 ATCTGATTAAGCTAGTGTGGGTAAAGGCTTCAAGGCGAGCATAGTATGTTGCTCTG 238  
QY 267 AGAGAGCGACGACCGACCTGAGGAGTGAAGACAGGCGCGAGCTCTTACGAGGAGCGACG 326  
DB 239 AAGAGAGCGACCGACCGACCTGAGGAGTGAAGACAGGCGCGAGCTCTTACGAGGAGCGACG 298  
QY 327 TGGGGAATTTTGGACATATGGGCGCAAGCTGATTCAGCAATGCGCGGTGATGAAGAAAG 386  
DB 299 TGGGGAATTTTGGACATATGGGCGCAAGCTGATTCAGCAATGCGCGGTGATGAAGAAAG 358  
QY 387 CCTTGGGCTTGAAGCTCTTCAAGTGAAGAAAGGTTACGTTAATTCGTGACT 446  
DB 359 CCTTGGGCTTGAAGCTCTTCAAGTGAAGAAAGGTTACGTTAATTCGTGACT 418  
QY 447 CATGACGCTATCGACAGAAAGACACCGGCTTAACTACGTCGACAGCGCGGTTAATACG 506  
DB 419 AATGACGCTATCGACAGAAAGACACCGGCTTAACTACGTCGACAGCGCGGTTAATACG 478  
QY 507 TAGGGTGCAGCGCTTAATCGGAATTAATCTGGGCGTTAAAGGCTGCGAGCGGCTTTGTAG 566  
DB 479 TAGGGTGCAGCGCTTAATCGGAATTAATCTGGGCGTTAAAGGCTGCGAGCGGCTTTGTAG 538  
QY 567 TCGATGCTGAATTCCTCGGCTTAACTCGGAGATGCGTTGAATACCAAGGCTAGAGT 626  
DB 539 TCGATGCTGAATTCCTCGGCTTAACTCGGAGATGCGTTGAATACCAAGGCTAGAGT 598  
QY 627 GTGGCAGAGGAGGTGGAATTCATGTGTAGCAGTGAATTCGCTAGATATGGAAGAAC 686  
DB 599 CTGGCAGAGGAGGTGGAATTCATGTGTAGCAGTGAATTCGCTAGATATGGAAGAAC 658  
QY 687 ATGATGCGCAGAGGCGACCTCTCTGGGTTAACTGACGCTCTATGACGAAAGCGTGGGGA 746  
DB 659 ACCGATGCGCAGAGGCGACCTCTCTGGGTTAACTGACGCTCTATGACGAAAGCGTGGGGA 718  
QY 747 GCAAACGAGATTGATACCTCGTAGTCAAGCGCTTAAAGATGCAACTAGTTGTGGG 806  
DB 719 GCAAACGAGATTGATACCTCGTAGTCAAGCGCTTAAAGATGCAACTAGTTGTGGG 778  
QY 807 CTTTATTAAGCTTGTATACGAGTAAAGCTGAGTGAATGACCGCTGAGGAGTACGCTG 866  
DB 779 TCTTATTAAGCTTGTATACGAGTAAAGCTGAGTGAATGACCGCTGAGGAGTACGCTG 838  
QY 867 CAAGATTAAACTCAAGAGATTGACGAGGAGCGCGCAAGCGGTGATTAATGAGATTA 926  
DB 839 CAAGATTAAACTCAAGAGATTGACGAGGAGCGCGCAAGCGGTGATTAATGAGATTA 898  
QY 927 ATTGATGCAACGCGGAAACCTTACCTTACCTTGAATGAGATTTTCTAGAGATA 986  
DB 899 ATTGATGCAACGCGGAAACCTTACCTTGAATGAGATTTTCTAGAGATA 958  
QY 987 GATTAGTGTCT--TCGGGAACGCTTAACACAGGTGCTGATGCTGCTGCTGCTGCTC 1043  
DB 959 GGGAGTGTCTGAAAGAGAACCACTATACAGGTGCTGATGCTGCTGCTGCTGCTGCTC 1018  
QY 1044 GTGAGATGTTGGTTAAGTCTCCGCAACGAGCGCAACCTTGTCAATAGTTCTC-----TAC 1103  
DB 1019 GTGAGATGTTGGTTAAGTCTCCGCAACGAGCGCAACCTTGTCAATAGTTCTC-----TAC 1073  
QY 1104 GGTGGGCACTTTATGAGATGCTGCGGTGAACAAACCGGAGGAGAGGTGGGAGTACGTTAA 1163  
DB 1074 GAAAGGGCACTCTATATGAGACTGCGGTGAACAAACCGGAGGAGAGGTGGGAGTACGTTAA 1133  
QY 1164 GTCTCATGCGCTTATGAGGTGAGGCTTCAACAGTATATCAATGCGCGTTCAGAGGGTT 1223

DB 1134 GTCTCATGCGCTTATGAGGTGAGGCTTCAACAGTATATCAATGATATACAGAGCGCC 1193  
QY 1224 GCCAACCCGCGAGGGGAGCTTAATCTCAAGAAAGCGCTGATGATCGGATCGGATCTGC 1283  
DB 1194 GCCAACCCGCGAGGGGAGCTTAATCGAAGAAAGTATATGATGTCGGAATGTATCTGTC 1253  
QY 1284 AACTGACTCTCGTGAATCGGAATGCTAGTAATGCGGATACGATGCGCGTGAATA 1343  
DB 1254 AACTGACTCTCGTGAATGCTAGTAATGCGGATACGATGCGCGTGAATA 1313  
QY 1344 GGTCCCGGGTCTTATACACACCGCGCTCAACCATGGAAGTGGTTTACCAAGACA 1403  
DB 1314 GGTCCCGGGTCTTATACACACCGCGCTCAACCATGGAAGGCGGTTTACCAAGAGTA 1373  
QY 1404 GGTAGCTTAACCGTAAGAGAGGCGGCTT 1430  
DB 1374 GGTAGCTTAACCGTAAGAGAGGCGGCTT 1400

RESULT 36  
ADE11077  
ID ADE11077 standard; DNA; 1481 BP.  
XX AC ADE11077;  
XX DT 29-JAN-2004 (first entry)  
XX DE Acidovorax avenae subsp. citrulli 16S ribosomal RNA gene SEQ ID NO:1.  
XX KM 16S-23S spacer DNA; bacterial; Acidovorax avenae subspecies citrulli;  
XX KW Xanthomonas cucurbitae; Erwinia tracheiphila; bacterial pathogen;  
XX KM 16S ribosomal RNA; gene; ds.  
XX OS Acidovorax avenae subsp. citrulli.  
XX EN WO2003082079-A2.  
XX PD 09-OCT-2003.  
XX PF 25-MAR-2003; 2003WO-US009618.  
XX PR 25-MAR-2002; 2002US-0367628P.  
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX PI Beck JÜ, Barnett CÜ, Yarnall MS, Zeilouni L;  
XX DR WPI; 2003-902900/82.  
XX PT New nucleic acid encoding a 16S-23S spacer DNA sequences for bacterial  
XX PT species Acidovorax avenae subspecies citrulli, Xanthomonas cucurbitae and  
XX PT Erwinia tracheiphila, useful in identifying A. avenae subspecies citrulli  
XX PT in melons.  
XX PS Example 8; SEQ ID NO 1; 50bp; English.  
XX CC The present invention describes a nucleic acid molecule encoding a 16S-  
XX CC 23S spacer DNA sequence for the bacterial species Acidovorax avenae  
XX CC subspecies citrulli or avenae, Xanthomonas cucurbitae and Erwinia  
XX CC tracheiphila. Also described: (1) a pair of oligonucleotide primers,  
XX CC where at least one primer comprises 15-18, 15-22 or 20-25 bp; (2)  
XX CC detecting bacterial pathogen; (3) a diagnostic kit used in detecting the  
XX CC bacterial pathogen; (4) a polypeptide comprising the amino acid sequence  
XX CC (S1, see ADE11075); (5) an antibody that reacts with a polypeptide having  
XX CC the N-terminal amino acid sequence of the polypeptide; and (6) an  
XX CC immunoassay for detecting Acidovorax avenae subspecies citrulli that uses  
XX CC the antibody. The nucleic acid is useful in diagnostic assays for  
XX CC identifying Acidovorax avenae subspecies citrulli in melons. The present  
XX CC sequence represents an Acidovorax avenae subspecies citrulli 16S  
XX CC ribosomal RNA gene, which is used in the exemplification of the present  
XX CC invention.

SQ Sequence 1481 BP; 378 A; 341 C; 470 G; 292 T; 0 U; 0 Other;  
 Query Match 77.5%; Score 1129.2; DB 10; Length 1481;  
 Best Local Similarity 87.3%; Pred. No. 0;  
 Matches 1275; Conservative 0; Mismatches 173; Indels 12; Gaps 3;  
 QY 1 ATTTGAAGCTGGCGGCGATCTTTACACATGACGAAAGTGGAGGAGCGAGATGCTTGCAAT 60  
 DB 1 ATTTGAAGCTGGCGGCGATCTTTACACATGACGAAAGTGGAGGAGCGATGCTTGCGA--- 57  
 QY 61 CTGTGTGGCAGTGGCGGAGACGGGTGAGTAAATGATCGGAACTGATCCAGAAAGAGGGGGAT 120  
 DB 58 -TGCTGACGAGTGGCGGAGAGGGGTGAGTAAATGATCGGAAAGTGGAGGAGGAGAT 116  
 QY 121 ACGCATCGAAAGATGTCTTAATACCGCATATACCTTAAGAGAGAAAGCGGGGATCGAAA 180  
 DB 117 ACGAGGCGGAAAGCTTTGCTAATACCGCATATAGATGATGAAAGCGGGGACCGGTAA 176  
 QY 181 GACCTTTGGCGCTTTGGAGGGCGGAGATGCTGATTTAGCTAGTTGGTGGGTAAAGGCTAC 240  
 DB 177 GGCCTTGGCGGAGACGAGGCGCGGATGCGAGATTAGTATGTTGGTGGGTAAAGGCTTAC 236  
 QY 241 CAAAGCGACATCAGTATGTTGTTCTGAGAGAGACGACCACTGAGGACTGAGACAGC 300  
 DB 237 CAAAGCTTACATCTGTAGCTGTGTGAGAGAGACGACCACTGAGGACTGAGACAGC 296  
 QY 301 GCCCAGACTCTTACCGGAGAGGACGACGTGGGGAATTTTGAACAATGGGCGCAAGCTGATC 360  
 DB 297 GCCCAGACTCTTACCGGAGAGGACGAGTGGGGAATTTTGAACAATGGGCGCAAGCTGATC 356  
 QY 361 CAGCAATGCGCGGTGAGTAAAGAGGCTTCGGGGTTGTAAGCTCTTTCACTCGAGAAAG 420  
 DB 357 CAGCAATGCGCGGTGAGTAAAGAGGCTTCGGGGTTGTAAGCTCTTTCACTCGAGAAAG 416  
 QY 421 AAAAGTTACGTAAATATCTGACTCATGACGATTCGACAGAAAGACACCGGCTAAC 480  
 DB 417 AAAAGCTTCTTCTAATAAAGGGGGCTCATGACGATTCGTAAGAAATACACCGGCTAAC 476  
 QY 481 TACGTGCGACAGCCGCGTAAATACGTAGGGTGCAAGCGTTAATCGGAATTTACTGGGCT 540  
 DB 477 TACGTGCGACAGCCGCGTAAATACGTAGGGTGCAAGCGTTAATCGGAATTTACTGGGCT 536  
 QY 541 AAAAGGTGGCGAGCGCGCTTTGTAAGTCAATGCTGAATCCCGGGCTTAACCTGGGAAT 600  
 DB 537 AAAAGGTGGCGAGCGCGTAAATACGTAGGGTGCAAGCGTTAATCGGAATTTACTGGGCT 596  
 QY 601 TCGCTTTGAAACTTACAAAGCTAGAGTGTGGCAGAGGAGAGTGGAAATTCATGTGTAGCAG 660  
 DB 597 TCGCATTTTGAAGTCAATGCTGGAATGACGAGAGGGGATGGAAATTCGCGGTGTAGCAG 656  
 QY 661 TGAATATGCGTAGATATGAAAGAAATGATGAGCGGAGCGACGCTCTGGGTTAACT 720  
 DB 657 TGAATATGCGTAGATATGCGAGAGAAACAGATGCGCAAGCAATCCCTGGGCTGTACT 716  
 QY 721 GACGCTCAACGAGAAAGGTGGGAGCAAAAGATTGATACCTGTTAGTCCACGCGC 780  
 DB 717 GACGCTCAACGAGAAAGGTGGGAGCAAAAGATTGATACCTGTTAGTCCACGCGC 776  
 QY 781 CTAAACGATGTCAACTAGTGTGTGGGCTTATTTAGGCTTGTATACGAACTAACCGGTGA 840  
 DB 777 CTAAACGATGTCAACTAGTGTGTGGGCTTCACTGATCAATGATACGAACTAACCGGTGA 836  
 QY 841 AGTTGACCGCTGGGAGATGAGGTGCGAAGATTTAAATCTCAAGGAATTGACGGGAGCC 900  
 DB 837 AGTTGACCGCTGGGAGATGAGGTGCGAAGATTTAAATCTCAAGGAATTGACGGGAGCC 896  
 QY 901 GACAAAGCGGTGATATGATGATTAATTCGATGCAACCGGAAAACCTTACCTACCTT 960  
 DB 897 GACAAAGCGGTGATATGATGATTTAATTCGATGCAACCGGAAAACCTTACCTACCTT 956  
 QY 961 GACATGTAGCAATTTTCTAGAGATAGATAGTCT---TCGGGAACGCTTAACACAGGTG 1017  
 DB 957 GACATGTAGCAATCTTTAGAGATAGAGAGTGTCTCGAAAGAAACCGTAAACACAGGTG 1016

QY 1018 CTGACATGCTGTGTCAGCTGTGTCGAGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1077  
 DB 1017 CTGACATGCTGTGTCAGCTGTGTCGAGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1076  
 QY 1078 ACCCTTGTCAATTAATTTGCGATCAATTTGGTGGGCACTTAAATGAGACTGCCGTGACAAA 1137  
 DB 1077 ACCCTTGTCAATTAATTTGCGATCAATTTGGTGGGCACTTAAATGAGACTGCCGTGACAAA 1131  
 QY 1138 CCGGAGGAAGTGGGAGATGAGTCAAGTCTCTATGCCCCCTTATGGGTAGGGCTTACACAG 1197  
 DB 1132 CCGGAGGAAGTGGGAGATGAGTCAAGTCTCTATGCCCCCTTATGGGTAGGGCTTACACAG 1191  
 QY 1198 TAAATCAATGGCGGTGACAGAGGGTGGCAACCGCGAGGGGGAGCTAATCTCAGAAAGC 1257  
 DB 1192 TCAATCAATGGCGGTGACAGAGGGTGGCAACCGCGAGGGGGAGCTAATCTCAGAAAGC 1251  
 QY 1258 GCGTGTAGTCCCGATCCGAGTCTGCAACTGACTCCGTGAAGTGGAAATGCTAGTAAT 1317  
 DB 1252 CAGTGTAGTCCCGATCCGAGTCTGCAACTGACTCCGTGAAGTGGAAATGCTAGTAAT 1311  
 QY 1318 CCGGATCAGCATGTGCGGTGATACGTTCCCGGCTTTGTACACACCCGCCGTACAC 1377  
 DB 1312 CCGGATCAGCATGTGCGGTGATACGTTCCCGGCTTTGTACACACCCGCCGTACAC 1371  
 QY 1378 CATGGAGTGGGTTTCAACAGAGCAGGTAGTCTAACCGTAAAGAGGCGCTTGGCACAG 1437  
 DB 1372 CATGGAGTGGGTTTTCGCAAGAGTGGTAACTTACCGTAAAGAGGCGCTTACACAG 1431  
 QY 1438 TGAGATTGATGACTGGGGTG 1457  
 DB 1432 CAGGTTGCTGACTGGGGTG 1451  
 RESULT 37  
 ID AB269298 standard; DNA; 1400 BP.  
 AC AB269298;  
 XX  
 DT 11-AUG-2003 (first entry)  
 XX  
 DE J lividum 16s ribosomal RNA gene fragment #2.  
 XX  
 KM Osteoarthritis; antibacterial; Janthinobacterium; 16s RNA; gene;  
 KW osteopathic; antiarthritic; analgesic; anti-inflammatory; de.  
 OS Janthinobacterium lividum.  
 XX  
 PN MO2002102384-A1.  
 XX  
 PD 27-DEC-2002.  
 XX  
 PF 17-JUN-2002; 2002WO-GB002771.  
 XX  
 PR 15-JUN-2001; 2001GB-00014672.  
 XX  
 PA (ORTH-) ORTHOGENICS AS.  
 PA (GARD/) GARDNER R.  
 XX  
 PI E1-Gewely MR;  
 XX  
 DR WPI; 2003-175199/17.  
 XX  
 PT Use of an antibacterial agent in the manufacture of a medicament for  
 PT treating osteoarthritis.  
 XX  
 PS Example 1; Page 46-48; 89pp; English.  
 CC The present invention relates to the use of an antibacterial agent in the  
 CC production of a treatment for osteoarthritis. The bacteria causing  
 CC osteoarthritis is Janthinobacterium lividum. The present sequence is a  
 CC fragment of the 16s RNA coding sequence from J. lividum shown in the

CC exemplification of the invention  
 XX Sequence 1400 BP; 371 A; 307 C; 432 G; 290 T; 0 U; 0 Other;  
 SQ

Query Match 77.4%; Score 1128.2; DB 8; Length 1400;  
 Best Local Similarity 89.1%; Pred. No. 0;  
 Matches 1354; Conservative 0; Mismatches 143; Indels 10; Gaps 3;

```

QY 27 CATGCAGTCGAAACGGCAGCAGATGCTTGATCTGTGTGGAGTGGCGAGCGGTGAG 86
DB 1 CATGTATAGTCGAAACGGCAGCAGCGA--GCTTGCTGTGTGTGGAGTGGCGAGCGGTGAG 58
QY 87 TAATGCATCGAACGATTCAGAAAGGCGGTGATACGATTCGAAAGATGTCTAATACCG 146
DB 59 TAATATATCGGAAACGATACCTAGAGTGGGGGATTAAGTACGAAAGTTACGCTTAATACCG 118
QY 147 CATATATCTTAAAGAGAGAAAGCAGGGGATCGAAAGCCTTGCGCTTTGAGCGCCGAT 206
DB 119 CATACGATCTTAAGAGTGAAGAGTGGGGATCGCAAGACCTCATGCTGTGGAGCGCCGAT 178
QY 207 GTCGATTAAGCTAGTGTGGGTAAAGGCTTACCAAGCGCAGATCAGTATGTTGCTCG 266
DB 179 ATCTGATTAAGCTAGTGTGGGTAAAGGCTTACCAAGCGCAGATCAGTATGTTGCTCG 238
QY 267 AGAGAGCAGACCACTGAGCACTGAGACACGCGCCAGACTCTTACCGGAGGACGACG 326
DB 239 AGAGAGCAGACCACTGAGCACTGAGACACGCGCTCAGACTCTTACCGGAGGACGACG 298
QY 327 TGGGGAAATTTTGAACAATGGGCGCAAGCCTGATTCAGCAATGGCGGGTGAAGAGAG 386
DB 299 TGGGGAAATTTTGAACAATGGGCGCAAGCCTGATTCAGCAATGGCGGGTGAAGAGAG 358
QY 387 CCTTGGGTTGTAAGCTCTTTCAGTCGAGAGAGAAAGTTACGTAATATATCTGACT 446
DB 359 CCTTGGGTTGTAAGCTCTTTCAGTCGAGAGAGAAAGTTACGTAATATATCTGACT 418
QY 447 CATGACGATTCAGCAAGAGAGACCGGCTTAACTAGTCCAGACCGCGGTAAATACG 506
DB 419 AATGACGATTCAGCAAGAGAGACCGGCTTAACTAGTCCAGACCGCGGTAAATACG 478
QY 507 TAGGGTGAAGCGTTAATCGGAATTAATCTGGGCGTAAAGGTTGGCGCGGCTTTGTAAG 566
DB 479 TAGGGTGAAGCGTTAATCGGAATTAATCTGGGCGTAAAGGTTGGCGCGGCTTTGTAAG 538
QY 567 TCAGATGTGAATCCCGGGGCTTAACTGGGAAATGCGTTGAAACTCAAGGCTAGAGT 626
DB 539 TCAGATGTGAATCCCGGGGCTTAACTGGGAAATGCGTTGGAATGGAATGGAAT 598
QY 627 GTGGCAGAGGAGGTGGAATTCATGTGTAGCAGTGAATGCGTGAATATGGAAGAC 686
DB 599 CTGGCAGAGGAGGTGGAATTCATGTGTAGCAGTGAATGCGTGAATATGGAAGAC 658
QY 687 ATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 746
DB 659 ACCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 718
QY 747 GCAAACGAGATTAATACCTGAGTACAGGCTTAAAGATGCACTAGTGTGAGG 806
DB 719 GCAAACGAGATTAATACCTGAGTACAGGCTTAAAGATGCACTAGTGTGAGG 778
QY 807 CCTATTATAGGCTTGTATACGAAAGCTAAACGCTGAAGTTGACCGCTGGGGAGTACGCTCG 866
DB 779 TCTTATATGACTTGTATACGAAAGCTAAACGCTGAAGTTGACCGCTGGGGAGTACGCTCG 838
QY 867 CAAGATTTAAACTCAAGAGATTAACGAGGAGCAGGAGCAGAGCGGTGATTAATGAGATTA 926
DB 839 CAAGATTTAAACTCAAGAGATTAACGAGGAGCAGGAGCAGAGCGGTGATTAATGAGATTA 898
QY 927 ATTGATTCGAACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 986
DB 899 ATTGATTCGAACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 958
QY 987 GATTATGTCCT--TCGGGAGCTTAACAGAGTGTGCTGCTCAGCTGCTGTC 1043

```

```

DB 959 AGGAGTCTCGAAAGAGAAACAGATACAGAGTGTGATGCTGTGCTGCTGCTGCTC 1018
QY 1044 GTGAGATGTTGGGTTAAGTCCCGCAACGAGGCAACCTTGTCAATTAATGCAATCATTT 1103
DB 1019 GTGAGATGTTGGGTTAAGTCCCGCAACGAGGCAACCTTGTCAATTAATGCTGCTGCTC 1073
QY 1104 GGTGGGCACTTAAATGAGTACGCGGTGACAAACCGGAGAGGAGTGGGATGACGTCA 1163
DB 1074 GAAAGGCACTTAAATGAGTACGCGGTGACAAACCGGAGAGGAGTGGGATGACGTCA 1133
QY 1164 GTCTCATGAGCCTTAAATGAGTACGCGGTGACAAACCGGAGAGGAGTGGGATGACGTCA 1223
DB 1134 GTCTCATGAGCCTTAAATGAGTACGCGGTGACAAACCGGAGAGGAGTGGGATGACGTCA 1193
QY 1224 GCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1283
DB 1194 GCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1253
QY 1284 AACTGACTCCGTTAAGTCCGAAATGCTAGTAAATGCGGATCAGCATGCTCGGTAATA 1343
DB 1254 AACTGACTCCGTTAAGTCCGAAATGCTAGTAAATGCGGATCAGCATGCTCGGTAATA 1313
QY 1344 CGTCCCGGAGTCTTGTACACACCGCGGTCAACCATGAGAGTGGGTTTACACAGAGCA 1403
DB 1314 CGTCCCGGAGTCTTGTACACACCGCGGTCAACCATGAGAGTGGGTTTACACAGAGCA 1373
QY 1404 GGTAGTCTAACCGTAAAGAGGCGCTT 1430
DB 1374 GGTAGTCTAACCGTAAAGAGGCGCTT 1400

```

RESULT 38  
 AD067895  
 ID AD067895 standard; DNA; 1509 BP.  
 XX  
 AC AD067895;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX  
 DE Acidovorax temperans partial 16S rRNA gene SEQ ID NO:2.  
 XX  
 KW de; gene; 16S rRNA; algae; algae-lysing bacteria; blue-green algae;  
 KW Anabaena cylindrica.  
 XX  
 OS Acidovorax temperans.  
 XX  
 PN KR2003075871-A.  
 XX  
 PD 26-SEP-2003.  
 XX  
 PF 21-MAR-2002; 2002KR-00015343.  
 XX  
 PR 21-MAR-2002; 2002KR-00015343.  
 XX  
 PA (HANY-) HANYANG HAK WON CO LTD.  
 XX  
 PI Bang SW, Choi JH, Han MS, Kim BH;  
 XX WPI; 2004-105157/11.  
 DR  
 XX  
 PT Algae-lysing bacteria Acidovorax temperans KfCC 11294 and method for  
 XX removing algae using the same microorganism.  
 XX  
 PS Example 3; SEQ ID NO 2; 1bp; Korean.  
 XX  
 CC The invention relates to a novel algae-lysing bacteria Acidovorax  
 CC temperans KfCC 11294 and a method for removing algae using the same  
 CC microorganism, thereby effectively solving environmental problems caused  
 CC by algae. An algae-lysing bacteria Acidovorax temperans KfCC 11294  
 CC containing 16S rDNA having the nucleotide sequence of AD067894 is  
 CC provided, wherein Acidovorax temperans KfCC 11294 is isolated from  
 CC Seokhon lake. A method for removing algae comprises culturing Acidovorax

temperans KFOC 11294 and applying the cultured medium to algae living in water, where the cultured medium of *Acidovorax* temperans KFOC 11294 is obtained by culturing *Acidovorax* temperans KFOC 11294 in BG-11 medium and centrifuging the cultured medium. The algae is blue-green algae, and the blue-green algae is *Anabaena cylindrica*. The present sequence represents the *Acidovorax* temperans partial 16S rRNA gene.

SQ Sequence 1509 BP; 389 A; 343 C; 477 G; 300 T; 0 U; 0 Other;

Query Match	77.48;	Score 1127.6;	DB 12;	Length 1509;
-------------	--------	---------------	--------	--------------

Best Local Similarity 87.3%; Pred. No. 0/

Matches 1274; Conservative 0; Mismatches 174; Indels 12; Gaps 3;

OY	1	ATTGAACGTGGCGCATCTTTACACATGCAAGTGAACGGCAGACCGAATGCTTCGAT	60
Db	19	ATTGAACGTGGCGCATCTTTACACATGCAAGTGAACGGTAAACGGTCTTCGGA----	75
OY	61	CTGTGGCGAGTGGCGGACGGGTGAAGTAATGCATCGGAACGTATCCGAAGAGGGGGTA	120
Db	76	-TGCTGACGAGTGGCGGAACGGGTGAGTAATATCATCGGAACGTGCCGATCTGTGGGGATA	134
OY	121	ACGCATCCGAABAATGCTTAATACCGCATATATCTTAAGAGGGAACGAGGGGATCGAAT	180
Db	135	ACGAAGCGAAGACTTTGCTTAATACCGCATTAAGATCTACGATGAAGAAGCAAGGGGACCGCA	194
OY	181	GACCTTGGCTTTTGGAGCGGCGCATGTCGTGATTAAGCTAGTTGGTGGGTAAAGGCTTAC	240
Db	195	GGCCTTGGCGGAACGGAGCGGCGCATGCGAGATTAAGTATGTTGGTGGGATTAAGGCTTAC	254
OY	241	CAGCGCAGCATCAAGTAGTTGGTCTGAGAGGACGACACACCTGGGACTGAGACAG	300
Db	255	CAGCGCAGCATCTTAAGCTGTGCTGAGAGGACGACACACACTGGGACTGAGACAG	314
OY	301	GCCCGACCTCTTAACGGGAGGACAGAGTGGGGAATTTTGACATGGGCGCCAAAGCTGATC	360
Db	315	GCCCGACCTCTTAACGGGAGGACAGAGTGGGGAATTTTGAACAATGGGCGAAGACCTGATC	374
OY	361	CAGCAATGCCCGGTGAGTAAGAAAGGCTTTCCGGTTGTAAAGCTCTTTCAGTGAGAGA	420
Db	375	CAGCATGCCCGGTGAAGGATAAAGGCTTTCCGGTTGTAAAGCTCTTTGTGTAACGAA	434
OY	421	AAAGCTTAACGGTAATTAATCGTGACTCATGACGGTATGACAGAAAGACACCGGCTTAC	480
Db	435	AAAGACTCTGGTTAATTAATCGTGGGCTCATGACGGTATCGTAAGATTAAGACACCGGCTTAC	494
OY	481	TACGTGCCAGCAGCCGCGGTAAATACGTAGGGTGCAGACGTTAATCGGAATTAATCGGGCGT	540
Db	495	TACGTGCCAGCAGCCGCGGTAAATACGTAGGGTGCAGACGTTAATCGGAATTAATCGGGCGT	554
OY	541	AAAGGTGGCGAGCGGCTTTGTAACTCAGATGTGAATATCCCGGGCTTAATCTGGGAAT	600
Db	555	AAAGGTGGCGAGCGGCTTATATTAAGACAGATGTGAATATCCCGGGCTCAACTGGGAAC	614
OY	601	TGCGTTTGAATCTAACAGGCTAGAGTGTGGCAGAGGAGAGTGGGAATTCATGTGTAGCAG	660
Db	615	TGCAATTTGACCTGTATAGCTAGAGTACGGCAGAGGGGGAATGGAAATTCGCGCGTATGACAG	674
OY	661	TGAATATGCGTAGAGATATGAAGAAACATCATGATGATGGCGAAGGCAAGCTCTCGGTAAACAT	720
Db	675	TGAATATGCGTAGATATATGCGAAGAAACACCGATGGCGAAGCAATCCCTCGGCGCTGTAACT	734
OY	721	GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAATACCTCGTATGTCACGCGC	780
Db	735	GACGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAATACCTCGTATGTCACGCGC	794
OY	781	CTAAACGATGTCACATAGTTGTGGGCTTAATTAAGCTTGTGAACGAACCTAACCGGTGA	840
Db	795	CTAAACGATGTCACATGTTGTGGGCTTCACTGACTCAGTAAACGAAGCTAACCGGTGA	854
OY	841	AGTTGACCGGCTTGGGGAGTACGGTGCAGAAATTAATACTCAAAGAAATTGACGGGGACCC	900
Db	855	AGTTGACCGGCTTGGGGAGTACGGGCGCAAGGTTGAATCTCAAAGAAATTGACGGGGACCC	914

QY	901	GCACAGCGGGGANTATGATGGATTAATTCAGTCAGCGCAAAAACCTTACTACCTT	960
Db	915	GCAACAGGGGAGATGATGGTTTAAATTCAGTACAGCGAAAACCTTACCACCTT	974
QY	961	GACATGTAGCCAAATTTCTAGAGATAGATTAGTCT---TCGGGACGCTAACACAGGTG	101
Db	975	GACATGTACGGAATCTTTAGAGATAGAGAGTGTCTCGAAGAGAGCCGTAAACAGGTG	103
QY	1018	CTGCATGGCTGTCTGACCTGTGTCTGTGAGATGTTGGTTAAATTCGCCAACGAGCGCA	107
Db	1035	CTGCATGGCTGTCTGACCTGTGTCTGTGAGATGTTGGTTAAATTCGCCAACGAGCGCA	109
QY	1078	ACCCCTTGCAATTAATGCAATCATTTGGTTGGGCACTTAAATAGACTGCGGATGACAA	113
Db	1095	ACCCCTTGCAATTAATGCTTACGAAA----GCGCACTTAATGGAATGCTGCCGATGACAA	114
QY	1138	CCGAGAGAGAGGTGGGATGACGTCAAGTCTCATATGACCCCTTAATGAGGCTTACACAG	119
Db	1150	CCGAGAGAGAGGTGGGATGACGTCAAGTCTCATATGACCCCTTAATGAGTGGGCTTACACAG	120
QY	1198	TAATTAACAATGCGCGCTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAAC	125
Db	1210	TCATTAACAATGCTGGTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCCCATTAAC	126
QY	1258	GCGTCGTATGTCGGATTCGGAGTCTGCACACTGCATCTCCGTGAAGTGGGAATCCCTGTAT	131
Db	1270	CAGTCGTATGTCGGATTCGGAGTCTGCACACTGCATCTCCGTGAAGTGGGAATCCCTGTAT	132
QY	1318	CGCGGATCAGCATGTCGCGGATGAAATACGTTCCCGGGTCTTGTACACACCGCCGTCACAC	137
Db	1330	CGCGGATCAGAAATGTCGCGGATGAAATACGTTCCCGGGTCTTGTACACACCGCCGTCACAC	138
QY	1378	CATGGAGATGGGATTACACAGAGCAGGTAGTTAAACCTAAGAGAGGCGCTTGGCCACGG	143
Db	1390	CATGGAGATGGGATTCTGCCAGAAAGTAGGCTTAACGTAAGAGAGGCGCTTACACACGG	144
QY	1438	TGAGATTCAATCACTGGGGTG 1457	
Db	1450	CAGGTTCTGTGACTGGGGTG 1469	
RESULT 39			
ADB61690			
ID	ADB61690	standard; DNA; 1488 BP.	
AC	ADB61690;		
XX			
DT	04-DEC-2003	(first entry)	
XX			
DE	16S rRNA of Burkholderia mallei DNA sequence.		
XX			
KM	enriching mRNA; high quality bacterial mRNA; bacterial gene expression;		
KM	poly-A tail; mRNA purification; oligo-dT capture;		
KM	prokaryote mRNA purification; bridging oligonucleotide; targeting region;		
KM	capable oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;		
KM	eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;		
KM	28S eukaryotic rRNA bridging oligonucleotide; ds.		
XX			
OS	Burkholderia mallei.		
XX			
PN	WO2003054162-A2.		
XX			
PD	03-JUL-2003.		
XX			
PF	19-DEC-2002; 2002MO-US041014.		
XX			
PR	20-DEC-2001; 2001US-00029397.		
XX			
PA	(AMBI-) AMBION INC.		
XX			
PI	Murphy GL, Whitley JP;		
XX			
DR	WPI; 2003-663255/62.		

XX Depleting or isolating targeted nucleic acids e.g. RNA, involves using a  
PT bridging oligonucleotide comprising bridging region and a targeting  
PT region complementary to a targeted nucleic acid, and a capture  
PT oligonucleotide.

Claim 4, Page 174, 208pp, English.

XX This invention relates to a novel method for isolating, depleting or  
XX separating a targeted nucleic acid, such as RNA, from a sample  
XX comprising targeted and non-targeted nucleic acids. It effects a way of  
XX enriching for non-targeted nucleic acids such as mRNAs. Isolating  
XX sufficient quantities of high quality bacterial mRNA is a demanding  
XX process which impedes analysis of bacterial gene expression in the  
XX presence of host cells. A small percentage of bacterial mRNAs may be poly  
XX -A tailed, but these are targeted for degradation and tend to be  
XX unstable. As a result, the commonly employed method for mRNA purification  
XX with eukaryotic cells, oligo-dT capture, is ineffective. The present  
XX invention provides an alternative, more suitable method for mRNA  
XX purification from prokaryotes. The method of the invention comprises the  
XX incubation of a sample with a bridging oligonucleotide (containing a  
XX targeting region) and subsequently incubating with a capture  
XX oligonucleotide allowing the isolation of the target from the sample. The  
XX method is useful for depleting or isolating targeted nucleic acid, for  
XX example RNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S  
XX or 18S or eukaryotic 28S rRNA, from a sample. The rRNA sequence may  
XX comprise any one of 64 fully defined sequences as given in the  
XX specification. The present sequence is that of a DNA sequence which  
XX represents the sequence of 16S rRNA of *Burkholderia mallei* related to the  
XX invention.

XX Sequence 1488 BP; 373 A; 344 C; 480 G; 291 T; 0 U; 0 Other;

Query Match 77.3%; Score 1126.8; DB 10; Length 1488;

Best Local Similarity 87.2%; Pred. No. 0;

Matches 1273; Conservative 0; Mismatches 177; Indels 10; Gaps 3;

```

QY 1 ATTGAACGCTGGCGGCGATGCTTACACATGCAAGTCGACGCGACGCGATGCTTGAT 60
DB 3 ATTGAACGCTGGCGGCGATGCTTACACATGCAAGTCGACGCGACGCG--GCTTCGCG 60
QY 61 CTGGTGGCGATGGCGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 61 CTGGTGGCGATGGCGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 61 CTGGTGGCGATGGCGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 61 CTGGTGGCGATGGCGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 121 ACCGATGAAAGATGCTCTAATCCGATATACCTTAAGGAGGAAAGGAGGAGTGA 180
DB 121 GCGCGGCGGAAAGCGGATTAATACCGATACGATCTGAGATGAAAGGAGGAGGAGCTTCG 180
QY 121 GACCTTGCCTTTTGGAGCGGCGATGCTGATGATGATGATGATGATGATGATGATGAT 240
DB 181 GACCTTGCCTTTTGGAGCGGCGATGCTGATGATGATGATGATGATGATGATGATGAT 240
QY 181 GACCTTGCCTTTTGGAGCGGCGATGCTGATGATGATGATGATGATGATGATGATGAT 240
DB 181 GACCTTGCCTTTTGGAGCGGCGATGCTGATGATGATGATGATGATGATGATGATGAT 240
QY 241 CAAAGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 CAAAGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 GCCCAAGATCTTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 301 GCCCAAGATCTTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 301 GCCCAAGATCTTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 301 GCCCAAGATCTTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 361 CAGCAATGCGCGTGAAGTGAAGAGGCGCTTGGGTTGTAAGCTCTTCACTGAGAGA 420
DB 361 CAGCAATGCGCGTGAAGTGAAGAGGCGCTTGGGTTGTAAGCTCTTCACTGAGAGA 420
QY 421 AAAGTTACGTTAAATTAATGCTGATCATGACCGTATGACAGAGAGAGCGGCTTAC 480
DB 421 AAATCATTTCTGCTAATATCCCGAGTGAATGACCGGATGAGAGAGAGAGAGAGAG 480
QY 481 TAGCTGCGAGAGCGCGGCTTAATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 481 TAGCTGCGAGAGCGCGGCTTAATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
```

```

QY 541 AAAGGTGCGAGCGCGGCTTTGTAAGTGAATGTAATCCCGGCGCTTAACCTGGAA 600
DB 541 AAAGGTGCGAGCGCGGCTTTGTAAGTGAATGTAATCCCGGCGCTTAACCTGGAA 600
QY 601 TGGCTTTGAAACTACAGAGCTTGAAGTGGCGAGAGGAGGTGAATTTCCATGTGAGAG 660
DB 601 TGCATTGGAGCTGCGACAGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660
QY 661 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 GAGCGCTATGACGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
DB 721 GAGCGCTATGACGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 781 CTAAACGATGATCACTAGTGTGTTGGGCTTATTAAGCTTTGATGATGATGATGATGAT 840
DB 781 CTAAACGATGATCACTAGTGTGTTGGGCTTATTAAGCTTTGATGATGATGATGATGAT 840
QY 841 AGTTGACCGCTGGGAGGAGTACCGTCCGCAAGATTAATACTCAAGAGGAGGAGGAGGAG 900
DB 841 AGTTGACCGCTGGGAGGAGTACCGTCCGCAAGATTAATACTCAAGAGGAGGAGGAGGAG 900
QY 901 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 GCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
DB 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
QY 1018 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
DB 1018 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
QY 1021 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1078 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
DB 1081 ACCCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135
QY 1138 CCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1197
DB 1136 CCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1195
QY 1198 TAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257
DB 1196 TAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1255
QY 1258 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
DB 1256 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1315
QY 1318 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377
DB 1316 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375
QY 1378 CATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1437
DB 1376 CATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1435
QY 1438 TGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
DB 1436 TGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455
```

RESULT 40

AD067894 ID AD067894 standard; DNA; 1509 BP.

AD067894;

09-SEP-2004 (first entry)



DE Acidovorax temperans partial 16S rRNA gene SEQ ID NO:1.  
KM ds; gene; 16S rRNA; algae; algae-lysing bacteria; blue-green algae;  
KM Anabaena cylindrica.  
OS Acidovorax temperans; KFCC 11294.  
PN KR2003075871-A.  
PD 26-SEP-2003.  
XX 21-MAR-2002; 2002KR-00015343.  
XX 21-MAR-2002; 2002KR-00015343.  
XX 21-MAR-2002; 2002KR-00015343.  
XX (HANY-) HANYANG HAK WON CO LTD.  
XX Bang SW, Choi JH, Han MS, Kim BH;  
XX WPI; 2004-105157/11.  
XX Algae-lysing bacteria Acidovorax temperans KFCC 11294 and method for  
PT removing algae using the same microorganism.  
XX Claim 1; SEQ ID NO 1; 1pp; Korean.  
XX The invention relates to a novel algae-lysing bacteria Acidovorax  
CC temperans KFCC 11294 and a method for removing algae using the same  
CC microorganism, thereby effectively solving environmental problems caused  
CC by algae. An algae-lysing bacteria Acidovorax temperans KFCC 11294  
CC containing 16S rDNA having the nucleotide sequence of AD067894 is  
CC provided, wherein Acidovorax temperans KFCC 11294 is isolated from  
CC Seochon Lake. A method for removing algae comprises culturing Acidovorax  
CC temperans KFCC 11294 and applying the cultured medium to algae living in  
CC water, where the cultured medium of Acidovorax temperans KFCC 11294 is  
CC obtained by culturing Acidovorax temperans KFCC 11294 in BG-11 medium and  
CC centrifuging the cultured medium. The algae is blue-green algae, and the  
CC blue-green algae is Anabaena cylindrica. The present sequence represents  
CC the Acidovorax temperans KFCC 11294 16S rRNA gene.  
XX  
SQ Sequence 1509 BP; 386 A; 347 C; 477 G; 299 T; 0 U; 0 Other;  
Query Match 77.3%; Score 1126; DB 12; Length 1509;  
Best Local Similarity 87.2%; Pred. No. 0;  
Matches 12/3; Conservative 0; Mismatches 175; Indels 12; Gaps 3;  
QY 1 ATTGAACCTGGCGGCGATCTTTACATGCAAGTCGAAACGCGACGCGATCTTGCAAT 60  
DB 19 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAAACGCGATCTTGCGA--- 75  
QY 61 CTGTGGGAGTGGGCGGACGGGTGATATGATCGGAACTGATCAGAAAGGGGGGTA 120  
DB 76 -TGTGACGAGTGGGCGGACGGGTGATATGATCGGAACTGATCGTGGGGGATG 134  
QY 121 ACGCATCGAAAGATGTGATATACCGCATATCTCTAAGAGGAAAGCAGGGATCGAAA 180  
DB 135 ACGAAGCGAAAGCTTTGCTAATACCGCATACGATCTACGAAAGCAGGGGACCGCAA 194  
QY 191 GACCTTGGCTTTTGGAGCGGCGCATGTCTGATTAAGCTTAAGTTGGTGGGTAAGGCTTAC 240  
DB 195 GCGCTTGGCGGAAACGGAAGCGGCGCATGTGAGATTAGTGTGGGTAAGAACTTAC 254  
QY 241 CAAAGCGATCATGATGTTGTTCTGAGAGACGACCAAGCCTTGGGACTGAGAACG 300  
DB 255 CAAAGCGATCATGATGTTGTTCTGAGAGACGACCAAGCCTTGGGACTGAGAACG 314  
QY 301 GCCGAGACTCTTACGAGGAGCAGCATGTGGGAAATTTTGGACAATGGGCGCAAGCTTGATC 360  
DB 315 GCCGAGACTCTTACGAGGAGCAGCATGTGGGAAATTTTGGACAATGGGCGCAAGCTTGATC 374  
QY 361 CAGCAATGCCGCGTGAAGTGAAGAGAGCCCTTGGGTTGTAAAGCTCTTTCACTGAGAA 420  
DB 375 CAGCAATGCCGCGTGAAGTGAAGAGAGCCCTTGGGTTGTAAAGCTCTTTCACTGAGAA 434

QY 421 AAAGGTTACGGTAATTAATCGNACTCATGACGGATCGACAGAAAGACACCGGCTAAC 480  
DB 435 AAAGACTCTGGTTATTAATCTGGGCTCATGACGGATCGTAAGATTAAGACACCGGCTAAC 494  
QY 481 TACGTCCAGCAGCGCGCGGTAAATAGTAGGGTGCAGCGGTTAATCGGAATTAAGTGGCGT 540  
DB 495 TACGTCCAGCAGCGCGCGGTAAATAGTAGGGTGCAGCGGTTAATCGGAATTAAGTGGCGT 554  
QY 541 AAAGGTTCCGACAGCGGCTTTGTAAGTCAATGTGAAATCCCGGGCTTAACTGGGAAT 600  
DB 555 AAAGGTTCCGACAGCGGCTTTGTAAGTCAATGTGAAATCCCGGGCTTAACTGGGAAT 614  
QY 601 TGCGTTTAAACTCAAGAGCTAGAGTGTGGGACGAGGAGGTGGAATTCATGATGAGCAG 660  
DB 615 TGCAATTTGTGACTGTATAGTACGAGTACGCGACAGGAGGAGATGGAATTCGCGTGTAGCAG 674  
QY 661 TGAATTCGTAGAGATATGAAAGACATCGATGCGGAGAGGAGCGCTCTGGGTTAACACT 720  
DB 675 TGAATTCGTAGATATGCGGAGGAAACACGATGCGGAGGAGCAATCCCTGGGCGTGTACT 734  
QY 721 GACGCTCATGACGAAACCGTGGGAGGACAAACAGGATTAATCCCTGTTAGTCCAGCC 780  
DB 735 GACGCTCATGACGAAACCGTGGGAGGACAAACAGGATTAATCCCTGTTAGTCCAGCC 794  
QY 781 CTAAACGATGTCACTAGTTGGGCTTATTAAGCTTGTGTAACGAACTAACGCGTGA 840  
DB 795 CTAAACGATGTCACTAGTTGGGCTTATTAAGCTTGTGTAACGAACTAACGCGTGA 854  
QY 841 AGTTGACCGCTGGGAGATACCGTGCAGAAATTAATCAAGAAATTAAGCGGAGACC 900  
DB 855 AGTTGACCGCTGGGAGATACCGTGCAGAAATTAATCAAGAAATTAAGCGGAGACC 914  
QY 901 GCAAGGCGGTGATTAATGATTAATTCATGACAGCGGAAATTCCTTAACCTT 960  
DB 915 GCAAGGCGGTGATTAATGATTAATTCATGACAGCGGAAATTCCTTAACCTT 974  
QY 961 GACATGTAGCGAAATTTTCTAGATAGATTAGTGTCT---TCGGGACGCTAACACAGGTG 1017  
DB 975 GACATGTAGCGAAATTTTCTAGATAGATTAGTGTCTCGAAAGAAACCGCTTAACACAGGTG 1034  
QY 1018 CTGCATGCTGTGCTGACGCTGTGTGATGATGTGGGTTAAGTCCCGCAACGAGCGCA 1077  
DB 1035 CTGCATGCTGTGCTGACGCTGTGTGATGATGTGGGTTAAGTCCCGCAACGAGCGCA 1094  
QY 1078 ACCCTTGTCAATTAATTCGATCAATTTGGTGGGCACTTAATGAGACTGCGGTCACAA 1137  
DB 1095 ACCCTTGTCAATTAATTCGATCAATTTGGTGGGCACTTAATGAGACTGCGGTCACAA 1149  
QY 1138 CCGGAGGAAAGGTGGGAGATGACGTCAAGTCTCATGGCCCTTAATGGGTTAGGGCTTACACG 1197  
DB 1150 CCGGAGGAAAGGTGGGAGATGACGTCAAGTCTCATGGCCCTTAATGGGTTAGGGCTTACACG 1209  
QY 1198 TAATACATGCGCGGTACAGAGGGTTGCAACCCGCGAGGGGAGCTTAATTCGAAAGC 1257  
DB 1210 TCATACATGCGCGGTACAGAGGGTTGCAACCCGCGAGGGGAGCTTAATTCGAAAGC 1269  
QY 1258 GCGTGTAGTCCGATGCGGATGCGCACTGCACTCGGTGAAGTCCGAAATGCTTAAT 1317  
DB 1270 CAGTGTAGTCCGATGCGGATGCGCACTGCACTCGGTGAAGTCCGAAATGCTTAAT 1329  
QY 1318 GCGGATGAGCATGCGCGGTGAATACGTTCCGGGGCTTGTACACACCGCGGTACAC 1377  
DB 1330 GCGGATGAGCATGCGCGGTGAATACGTTCCGGGGCTTGTACACACCGCGGTACAC 1389  
QY 1378 CATGGAAGTGGGTTTCAACAGAACAGGTACTTAATCCGTAAGAGAGGGCGTTGCCACGG 1437  
DB 1390 CATGGAAGTGGGTTTCTGCAAGAAATGATGCTTAACGTTAAGAGGGCGTTTCAACGG 1449  
QY 1438 TGAATTCATGACTGGGGTG 1457  
DB 1450 CAGGTTTCGTGACTGGGGTG 1469



RESULT 41  
ADCS3929  
ID ADCS3929 standard; DNA; 1463 BP.  
XX  
XX ADCS3929;  
AC  
XX 18-DEC-2003 (first entry)  
XX  
XX  
XX Phenyl hydroxylase gene fragment #SEQ ID 1.  
DE  
XX  
XX Gene fixing; ecosystem; phenyl hydroxylase gene; phenol decomposition;  
KM sludge; purification; ds.  
XX  
XX  
XX Comamonas sp.  
XX JP200314165-A.  
XX  
XX  
XX 20-MAY-2003.  
XX  
XX 12-NOV-2001; 2001JP-00346473.  
XX  
XX 12-NOV-2001; 2001JP-00346473.  
XX  
XX (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.  
XX  
XX WPI; 2003-818164/77.  
XX  
XX  
XX Fixing useful genes e.g. phenol hydroxylase to micro organism ecosystems  
PT e.g. phenol decomposition activated sludge, involve isolating dominant  
PT strain in ecosystem, introducing a useful gene and returning the strain  
XX to the ecosystem.  
XX  
XX Example 2; SEQ ID NO 1; 20pp; Japanese.  
XX  
XX The invention relates to a method for fixing a useful gene in  
XX microorganisms ecosystems, comprising isolating a strain from  
XX microorganisms ecosystem, which is a dominant strain within the  
XX ecosystem, introducing a useful gene into the isolated strain and then  
XX returning the strain to the microorganisms ecosystem. The method is  
XX useful for fixing a useful gene e.g., phenyl hydroxylase gene to a strain  
XX dominant within a ecosystem, such as Comamonas sp. IN7 strain and its  
XX closely related strains. It is also useful for increasing phenol  
XX decomposition activity in a sludge, and for purifying environment  
XX contaminated with phenol. The current sequence represents a phenyl  
XX hydroxylase gene fragment.  
XX  
XX Sequence 1463 BP; 374 A; 344 C; 465 G; 280 T; 0 U; 0 Other;  
SQ  
Query Match 77.2%; Score 1124.8; DB 10; Length 1463;  
Best Local Similarity 87.2%; Pred. No. 0;  
Matches 1271; Conservative 0; Mismatches 177; Indels 10; Gaps 3;  
QY 3 TGAACGCTGGCGGCGATGCTTTACACATGCAAGTGAACGGGACGAGCTTGCACT 62  
DB 1 TGAACGCTGGCGGCGATGCTTTACACATGCAAGTGAACGGGACGAGCTTGCACT 58  
QY 63 GGTGGCGAGTGGCGGAGCGGGAGTATGATGCAAGTATCCAGAAAGAGGAGGATGCA 122  
DB 59 GATGGCGAGTGGCGGAGCGGGAGTATGATGCAAGTATCCAGAAAGAGGAGGATGCA 118  
QY 123 GCATCGAAAGATGTCTAATACCGCATATCTTAAGAGGAGAAAGAGGAGGATGCA 182  
DB 119 TACTCGAAAGATGTCTAATACCGCATATCTTAAGAGGAGAAAGAGGAGGATGCA 178  
QY 183 CTTTGGCTTTTGAAGCGCGGATGTCTGATTAAGTATGTTGGGGTAAAGGCTTACCA 242  
DB 179 CTTGGCGCTACCAAGCGGCTGATGCGAGTATGATGTTGGGGTAAAGGCTTACCA 238  
QY 243 AGGAGCGATGATGATGTTGTTGCTGAGAGGAGGACGAGCAGACCTGGGACTGAGA 302  
DB 239 AGCGGAGATCTGTAGTGTGTGTGAGAGGAGGACGAGCAGACCTGGGACTGAGA 298

QY 303 CCAGACTCTTAACGAGAGGAGCGATGGGAAATTTTGAACATATGGCGGACCTGATCCA 362  
DB 299 CCAGACTCTTAACGAGAGGAGCGATGGGAAATTTTGAACATATGGCGGACCTGATCCA 358  
QY 363 GCAATGCCGCTGAGTGAAGAAAGGCTTGGGTTGTAAGCTTTTCACTCGAAGAA 422  
DB 359 GCAATGCCGCTGAGTGAAGAAAGGCTTGGGTTGTAAGCTTTTCACTCGAAGAA 418  
QY 423 AGGTTACGTTAAATTAATCGTGAATCAATGACGATGACAGAAAGACCGGCTA 482  
DB 419 AGGCTTCTTAATTAACAGAGCCGATGACGATGACGATGACGATGACGATGAC 478  
QY 483 CGTCCAGAGCGCGCTGATTAATGATGAGTGAAGGCTTAATCGAATTAATCGAG 542  
DB 479 CGTCCAGAGCGCGCTGATTAATGATGAGTGAAGGCTTAATCGAATTAATCGAG 538  
QY 543 AGGTTGCGAGCGCGCTTGTGAATGATGATGAATTCCTCGGCTTAATCGGAA 602  
DB 539 AGGTTGCGAGCGCGCTTGTGAATGATGATGAATTCCTCGGCTTAATCGGAA 598  
QY 603 CGTTTGAATCTACAGGCTGATGATGATGATGATGATGATGATGATGATGAT 662  
DB 599 CTTTGTGACTGCAAGGCTGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 658  
QY 663 AATGCGTGAAGATATGAGAAACATGATGCGGAGGAGGAGGAGGAGGAGGAGGAG 722  
DB 659 AATGCGTGAAGATATGAGAAACATGATGCGGAGGAGGAGGAGGAGGAGGAGGAG 718  
QY 723 CGCTATGACAGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782  
DB 719 CGCTATGACAGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 778  
QY 783 AATGATGATCACTAGTTGTTGGGCTTATGAGCTTGTGAAGGCTTAACCGGTGA 842  
DB 779 AATGATGATCACTAGTTGTTGGGCTTATGAGCTTGTGAAGGCTTAACCGGTGA 838  
QY 843 TTGACCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 902  
DB 839 TTGACCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 898  
QY 903 ACAAGCGGTGATTAATGATGATTAATGATGATGATGATGATGATGATGATGAT 962  
DB 899 ACAAGCGGTGATTAATGATGATTAATGATGATGATGATGATGATGATGATGAT 958  
QY 963 CATGTAACCAATTTTCAAGATGATGATGATGATGATGATGATGATGATGAT 1019  
DB 959 CATGTAACCAATTTTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1018  
QY 1020 GCATGGCTGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079  
DB 1019 GCATGGCTGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078  
QY 1080 CTTTGTCAATTAATTTGCAATGATTTGTTGGGCACTTTAATGAGACTGCGGATG 1139  
DB 1079 CTTTGTCAATTAATTTGCAATGATTTGTTGGGCACTTTAATGAGACTGCGGATG 1133  
QY 1140 GAGAGAAAGTGGGAGTACGTAATGCTTCAATGAGCTTATGAGGATGCTTCA 1199  
DB 1134 GAGAGAAAGTGGGAGTACGTAATGCTTCAATGAGCTTATGAGGATGCTTCA 1193  
QY 1200 ATACAAATGAGGCGGTACAGAGGTTGCGCAACCGGAGGAGGAGGAGGAGGAG 1259  
DB 1194 ATACAAATGAGGCGGTACAGAGGTTGCGCAACCGGAGGAGGAGGAGGAGGAG 1253  
QY 1260 GTGTAATCCGAGATCGGAGTCTGCAATCTGCACTCGGTAAGTGGAAATCGCTA 1319  
DB 1254 GTGTAATCCGAGATCGGAGTCTGCAATCTGCACTCGGTAAGTGGAAATCGCTA 1313  
QY 1320 CGATGACGATGTCGCGGTGAATTCGTTCCGGGCTTTGATACACACCGCCCTG 1379  
DB 1314 TGATTCAGAAATGTCACGATGAATTCGTTCCGGGCTTTGATACACACCGCCCTG 1373  
QY 1380 TGGAGTGGGTTTCAACCAAGAGGATGATGATGATGATGATGATGATGATGAT 1439

Db 1374 TGGGAGCGGCTCGCAGAAAGTAGTACCTTAACCGCAAGAGGCGCTTACACAGGCG 1433  
Qy 1440 AGATTCATGACTGGGGTG 1457  
Db 1434 GGGTTCGTGACTGGGGTG 1451

RESULT 42  
ABZ69304  
ID ABZ69304 standard; DNA; 1482 BP.  
XX ABZ69304;  
AC  
XX 11-AUG-2003 (first entry)  
XX  
XX Jantchinobacterium 16S RNA coding sequence.  
DE  
XX Jantchinobacterium 16S RNA coding sequence.  
XX  
XX Osteoarthritis; antibacterial; Jantchinobacterium; 16S RNA; gene;  
KM osteopathic; antiarthritic; analgesic; anti-inflammatory; ds.  
XX  
XX Jantchinobacterium sp.  
OS  
XX MO2002102384-A1.  
PN  
XX 27-DEC-2002.  
PD  
XX 17-JUN-2002; 2002WO-GB002771.  
PF  
XX 15-JUN-2001; 2001GB-00014672.  
PR  
XX (ORTH-) ORTHOGENICS AS.  
PA (GARD/) GARDNER R.  
XX  
XX El-Gewely MR;  
PI  
XX WPI; 2003-175159/17.  
DR  
XX Use of an antibacterial agent in the manufacture of a medicament for  
PT treating osteoarthritis.  
PT  
XX Example 1; Fig 12; 89pp; English.  
PS  
XX The present invention relates to the use of an antibacterial agent in the  
CC production of a treatment for osteoarthritis. The bacteria causing  
CC osteoarthritis is Jantchinobacterium lividum. The present sequence is a  
CC fragment of the 16S RNA coding sequence from J. lividum shown in the  
CC exemplification of the invention

SQ Sequence 1482 BP; 381 A; 324 C; 439 G; 300 T; 0 U; 38 Other;  
Query Match 77.2%; Score 1124.6; DB 8; Length 1482;  
Best Local Similarity 88.3%; Pred. No. 0;  
Matches 1254; Conservative 0; Mismatches 156; Indels 10; Gaps 3;

Qy 27 CATGCAAGTCGAAACGGCAGACGCGATGCTTGCACTCTGCTGGCGAGTGGCGGACGGGTAG 86  
Db 44 CATGCAAGTCGAAACGGCAGACGCGA--GCTTGCTCTGCTGGCGAGTGGCGAAGGGGTAG 101  
Qy 87 TAAATGATCGGAACGTAATCCAGAAAGAGGGGGGTAACGATCGAAAGATGCTAATACG 146  
Db 102 TAAATATATCGAAGCTAACCTTAGATGGGGGATACGTAGCGAAAGTTACGCTAATACG 161  
Qy 147 CATATATCTTAAGAGAAAGCAGGGGATCGAAAGACCTTGCGCTTTTGGAGCGCGAT 206  
Db 162 CATAGCATCTTAAGATGAAGAGTGGGGATCGCAAGACCTCATGCTGTGGAAGGCGCGAT 221  
Qy 207 GTCTGATTGCTAAGTTGGTGGGTAAAGGCTTACCAAGCGAGATCAGTAAGTTGGTCTG 266  
Db 222 ATCTGATTGCTAAGTTGGTAAAGCTTACCAAGGATCGATCAGTAAGTTGGTCTG 281  
Qy 267 AAGAGGACGACGACGACGCTGAGACAGGCGCCAGACTCTACGGGAGGCGAG 326

Db 282 AAGAGACGACGACGACCTGGAATGAGACACGCTCCAGACTCTTACGGGAGCGAG 341  
Qy 327 TGGGGAATTTTGGACATATGGGCGCAAGCTTATCCAGCAATGCCGCTGAGTGAAGAG 386  
Db 342 TGGGGAATTTTGGACATATGGGCGCAAGCTTATCCAGCAATGCCGCTGAGTGAAGAG 401  
Qy 387 CCTTCGGGTTGTAAGCTCTTTCACTGAGTGAAGAAAGAAATGACGTAATATATCTGACT 446  
Db 402 CTTTCGGGTTGTAAGCTCTTTTGTCAAGGAAGAAACGCTGAGACTAATATCTTCTGCT 461  
Qy 447 CATGACGCTATCGACAGAAAGCAGCGCTAATCTAGCTGCCAGCAGCGCGCTAATACG 506  
Db 462 AATGACGCTGACTGAAGATTAAGCACCGCTTACTAGCTGCCAGCAGCGCGCTAATACG 521  
Qy 507 TAGGGTCAAGCGTTAATCGGAATTAATCGGCGTAAAGGTTGCGGAGCGGCTTTGTAG 566  
Db 522 TAGGGTCAAGCGTTAATCGGAATTAATCGGCGTAAAGGTTGCGGAGCGGCTTTGTAG 581  
Qy 567 TCAGATGTGAATATCCCGGGCTTAACTGGGAATTTGCGTTGAACCTAAGGCTAGAGT 626  
Db 582 TCTGATGTGAATATCCCGGGCTTAACTGGGAATTTGCGTTGAGACTGCAAGGCTAGAGT 641  
Qy 627 GTGCGAGAGGAGTGAATTCATGTGTAGCAGTGAATGCGTAGAATATGAAGAAC 686  
Db 642 CTGCGAGAGGAGGAGTGAATTCATGTGTAGCAGTGAATGCGTAGAATATGAAGAAC 701  
Qy 687 ATCGATGCGGAAGGAGGAGCTCTGGGTAAACCTGACCGCTCATGCGAAGGCGTGGGA 746  
Db 702 ACCGATGCGGAAGGAGGAGCTCTGGGTAAACCTGACCGCTCATGCGAAGGCGTGGGA 761  
Qy 747 GCAACAGGATTAAGATCCCTGTGTAGTCCAGCGCTTAAACGATGTCAACTAGTTGTGG 806  
Db 762 GCAACAGGATTAAGATCCCTGTGTAGTCCAGCGCTTAAACGATGTCTAGTTGTGG 821  
Qy 807 CCTTAATGAGCTTGTGTAAAGAACTAACGCTGAAAGTTGACCGCTGCGGAGTACGCTG 866  
Db 822 TCTTAATGACTTGTGTAAAGAACTAACGCTGAAAGTTGACCGCTGCGGAGTACGCTG 881  
Qy 867 CAAGTTTAAACTCAAGAAATTTGACGGGAGCCGCAAGGCGTGAATTTATGAGATTA 926  
Db 882 CAAGTTTAAACTCAAGAAATTTGACGGGAGCCGCAAGGCGTGAATTTATGAGATTA 941  
Qy 927 ATTGATGCAACGCGCAAAAACCTTACTACCTTGATGACATGTAGCAATTTTCTAGATA 986  
Db 942 ATTGATGCAACGCGCAAAAACCTTACTACCTTGATGACATGTGATGATTTCTTGAAGATC 1001  
Qy 987 GATTAGTCT--TCGGGAACGCTTAACAGGTGCTGATGCTGTGCTGCTGCTGCTG 1043  
Db 1002 AGGAGTGTCTGAAAGAAAGAACAGATACAGGTGCTGATGCTGTGCTGCTGCTGCTG 1061  
Qy 1044 GTGAATGTTGGGTTAATGCTCCGCAACGAGCGCAACCTTGCTAATTTATGCAATCATTT 1103  
Db 1062 GTGAATGTTGGGTTAATGCTCCGCAACGAGCGCAACCTTGCTAATTTATGCTGCTGCTG 1116  
Qy 1104 GGTGGGCACTTAATGAGACTGCGGTGACAAACCGAGAGAAAGGTGGGATGAGCTCA 1163  
Db 1117 GAAAGGCACTTAATGAGACTGCGGTGACAAACCGAGAGAAAGGTGGGATGAGCTCA 1176  
Qy 1164 GTCTCATGAGCCCTTAATGAGGCTTACACGTAATCAATATGCGCGCTACAGAGGTT 1223  
Db 1177 GTCTCATGAGCCCTTAATGAGGCTTACACGTAATCAATATGCTACATACAGAGCGCT 1236  
Qy 1224 GCGAACCGCGAGGGGAGCTAATCTCAGAAAGCCGCTGTATGTCGGAATGGAGTGGC 1283  
Db 1237 GCGAACCGCGAGGGGAGCTAATGCGAAAGTATCTGATGTCGGAATGTAATGCTGCTG 1296  
Qy 1284 AACTGACTCCGTGAAGTCGGAATGCTAATGCTGCGATACAGCATGTGCGGCTGAATA 1343  
Db 1297 AACTGACTGCTAAGTTGGAATGCTAATGCTGCGATACAGCATGTGCGGCTGAATA 1356  
Qy 1344 CGTTCCGGGCTTTTATACACACGCGCTGACACATGGAAGTGGGCTTTACCAAGAGCA 1403  
Db 1357 CGTTCCGGGCTTTTATACACACGCGCTGACACATGGAAGTGGGCTTTTACCAAGAGTA 1416



```
QY 1075 GCAACCTTGCATTAATTCATCATTTGTTGGGCACTTTAATGAGACGCGGTCAC 1134
DB 1109 GCAACCTTGCATTAATTCATCATTTGTTGGGCACTTTAATGAGACGCGGTCAC 1168
QY 1135 AAACCGGAGGAGTGGGGATGACGTCAGTCTCATGAGCCCTTAATGGTAAAGGCTTCA 1194
DB 1169 AAGCCGGAGGAGTGGGGATGACGTCAGTCTCATGAGCCCTTAATGAGAGGCTTCA 1228
QY 1195 ACGTAAATCAATGCGCGCTGACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAA 1254
DB 1229 ACGTAAATCAATGCGCGCTGACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAA 1288
QY 1255 AGCGGCTGCTACTCCGGATGCGAGTCTGCACTCGAATCGGAATCGGAATCGCTAGT 1314
DB 1289 AACCAGTCTGATCCGGATGCACTCTGCACTCGAATCGGAATCGGAATCGCTAGT 1348
QY 1315 AATCGCGATGACAGTATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTC 1374
DB 1349 AATCGCGATGACAGTATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTC 1408
QY 1375 CACCATGGAGTGGGTTTCAACCAAGACGTAATCTAACCGTAAGAGAGGCGCTTGCCA 1434
DB 1409 CACCATGGAGTGGGTTTCAACCAAGACGTAATCTAACCGTAAGAGAGGCGCTTGCCA 1468
QY 1435 CGGTGAGATTCAATGACTGGGGTG 1457
DB 1469 CGGTGAGATTCAATGACTGGGGTG 1491
```

## RESULT 44

```
AAA81490.00
WP Sequence split into 15 fragments LOCUS AAA81490 Accession Aaa81490
WP Fragment Name Begin End
```

```
WP AAA81490_00 1 110000
WP AAA81490_01 100001 210000
WP AAA81490_02 200001 310000
WP AAA81490_03 300001 410000
WP AAA81490_04 400001 510000
WP AAA81490_05 500001 610000
WP AAA81490_06 600001 710000
WP AAA81490_07 700001 810000
WP AAA81490_08 800001 910000
WP AAA81490_09 900001 1010000
WP AAA81490_10 1000001 1110000
WP AAA81490_11 1100001 1210000
WP AAA81490_12 1200001 1310000
WP AAA81490_13 1300001 1410000
WP AAA81490_14 1400001 1437668
ID AAA81490 Standard; DNA; 1437668 BP.
```

AAA81490;

04-DEC-2000 (first entry)

N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.

XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;  
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;  
KW Meningococcus B; MenB; ds.

OS Neisseria meningitidis.

PN WO200022430-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US023573.

PR 09-OCT-1998; 98US-0103794P.  
PR 30-APR-1999; 99US-0132068P.

XX (CHIR ) CHIRON CORP.

XX Frazer CM, Hickey B, Peterson J, Tettelin H, Venter JC;  
PI Maizumi V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
PI Rappuoli R, Pizza M;  
DR WPI, 2000-318079/27.  
PT Isolated nucleotide sequences of Neisseria meningitidis which can be used  
PT in the diagnosis and treatment of N. meningitidis infection and other  
PT Neisseria infections, for example, N.gonorrhoea.

Claim 7, Page 866-1272; 1760p; English.

CC The present invention describes methods of obtaining immunogenic proteins  
CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent  
CC specifically claimed Neisseria meningitidis genomic DNA sequences;  
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA  
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and  
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of  
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent  
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all  
CC used in the exemplification of the present invention. The nucleic acid  
CC sequences, protein sequences, and antibodies against them, can be used in  
CC the manufacture of a composition. The composition can be used as a  
CC medicament (or in the manufacture of a medicament) for treating,  
CC preventing or diagnosing infection due to Neisseria bacteria. For  
CC example, some of the identified proteins could be components of vaccines  
CC against Meningococcus B; against all serotypes; and/or against all  
CC pathogenic Neisseriae. Identification of sequences from the bacterium  
CC will also facilitate production of biological probes, particularly  
CC organism-specific probes. Attempts to make efficacious Meningococcus B  
CC vaccines have failed mainly due to antigen tolerance. Multivalent  
CC vaccines have also been tried but none have successfully overcome  
CC antigenic variability. The provision of further, complete sequences may  
CC provide an opportunity to identify secreted or surface exposed proteins  
CC that may be presumed targets for the immune system and which are not  
CC antigenically variable or at least more conserved than other more  
CC variable regions

SQ Sequence 1437668 BP; 344338A; 353206C; 385074G; 355045T; 0U; 50ther;

Query Match 77.2%; Score 1124.6; DB 3; Length 110000;

Best Local Similarity 87.4%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 179; Indels 6; Gaps 4;

1 ATTGAACCTGCGCGGATGCTTTACATGCAATGCAAGCGGACACAC--GGATGCTTGC 58

DB 60992 ATTGAACCTGCGCGGATGCTTTACATGCAATGCAAGCGGACACACAGAAAGCTTGC 61051

QY 59 ATCT--GGTGGCGAATGCGCGGATGCTTTACATGCAATGCAAGCGGACACAGAAAGGGG 116

DB 61052 TTCTCGGGTGGCGAATGCGCGGATGCTTTACATGCAATGCAAGCGGACACAGAAAGGGG 61111

QY 117 GGTAAAGCATGAAGATGCTTAATACCGCATATCTCTAAGAGAGAAAGAGGGGATC 176

DB 61112 GATACTGATCGAAAGATCAGCTTAATACCGCATATCTCTAAGAGAGAAAGAGGGGATC 61171

QY 177 GAAAGACCTTGGCGCTTTTGAAGCGGCGATGCTTAATGAGTGTGGGGGTAAAGGC 236

DB 61172 TTCTCGGGTGGCGAATGCGCGGATGCTTTAATGAGTGTGGGGGTAAAGGC 61231

QY 237 CTACCAAGGCGAGATCAGTATGTTGTTGAGAGACGACACGACCACTGGAGCTGAGA 296

DB 61232 CTACCAAGGCGAGATCAGTATGTTGTTGAGAGACGACACGACCACTGGAGCTGAGA 61291

QY 297 CAGGCGCCAGACTCTTACGGGAGGCGAGCAATGTTGGACAAATGGGCGCAAGCT 356

DB 61292 CAGGCGCCAGACTCTTACGGGAGGCGAGCAATGTTGGACAAATGGGCGCAAGCT 61351

QY 357 GATCCAGCAATGCGCGGATGAGTGAAGAGGCTTGGGTTGTTGAAGCTCTTCACTGAG 416

DB 61352 GATCCAGCAATGCGCGGATGAGTGAAGAGGCTTGGGTTGTTGAAGACTTTTTCAGGG 61411

417 AAGAAAGTTACGGTAAATATCTGATCTATGACGGTATGACAGAAAGACCGGC 476  
 61412 AAGAAAGGCTGTCTATATATCAGCGGTATGACGGTAACTGAAAGATTAAGACCGGC 61471  
 477 TAACTAGTGCAGACGCGCGGTAAATAGTAGGTGCAAGGGTTAAATCGGAATTACGG 536  
 61472 TAACTAGTGCAGACGCGCGGTAAATAGTAGGTGCAAGGGTTAAATCGGAATTACGG 61531  
 537 GCGTAAAGGAGTGCAGCGCGGTAAATAGTAGGTGCAAGGGTTAAATCGGAATTACGG 596  
 61532 GCGTAAAGGAGTGCAGCGCGGTAAATAGTAGGTGCAAGGGTTAAATCGGAATTACGG 61591  
 597 GAATTCGCTTTGAACTACAGGCTTAAAGTGTGCGAGAGGAGTGAATTCATGTGTA 656  
 61592 GAATTCGCTTTGAACTGCGGTGCTGAGTGTGCGAGAGGAGTGAATTCATGTGTA 61651  
 657 GCAATGAATTCGCTGAGATATGGAAGAACTGATGCGGAGGAGGAGTGAATTCATGTGTA 716  
 61652 GCAATGAATTCGCTGAGATATGGAAGAACTGATGCGGAGGAGGAGTGAATTCATGTGTA 61711  
 717 CACTAGCGCTATGACGAAAGGAGTGGAGCAAGAGATTAGATACCTGGTATGCA 776  
 61712 CACTAGCGCTATGACGAAAGGAGTGGAGCAAGAGATTAGATACCTGGTATGCA 61771  
 777 CCGCTTAAAGATGTCATAGTTGTTGGCCCTTATTA-GGCTTGTAAAGAACTTACG 835  
 61772 CCGCTTAAAGATGTCATAGTTGTTGGCCCTTATTA-GGCTTGTAAAGAACTTACG 61831  
 836 CGTGAAGTTGACCGCTGGGAGTACCGTTCGCAAGTTAAATCTCAAGAAATTGACGG 895  
 61832 CGTGAAGTTGACCGCTGGGAGTACCGTTCGCAAGTTAAATCTCAAGAAATTGACGG 61891  
 896 GACCCGCAAGAGGAGTGAATATGATGATTAATTCATGCAAGCGGAAATCTTACCTA 955  
 61892 GACCCGCAAGAGGAGTGAATATGATGATTAATTCATGCAAGCGGAAATCTTACCTA 61951  
 956 CCCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014  
 61952 GCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 62011  
 1015 GTGCTCATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074  
 62012 GTGCTCATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 62071  
 1075 GCAACCTTGTCTATTAATGCAATGCTTGTGGGCACTTATGAGACTGCGCGGTAC 1134  
 62072 GCAACCTTGTCTATTAATGCAATGCTTGTGGGCACTTATGAGACTGCGCGGTAC 62131  
 1135 AAACCGGAGAGGAGTGGAGTGAAGTCAAGTCTCATGAGCCCTTATGAGGCTTAC 1194  
 62132 AAACCGGAGAGGAGTGGAGTGAAGTCAAGTCTCATGAGCCCTTATGAGGCTTAC 62191  
 1195 ACCTAATACATGAGCGGTAAGAGGAGTGGCAAGCGGAGGAGGAGGAGGAGGAGGAG 1254  
 62192 ACCTAATACATGAGCGGTAAGAGGAGTGGCAAGCGGAGGAGGAGGAGGAGGAGGAG 62251  
 1255 AGGCGCTGATGATGCTGAGTCTGCAATCTGATCTGCGTGAAGTGGAGTGGAGT 1314  
 62252 AACCGATCTGATGCTGAGTCTGCAATCTGATCTGCGTGAAGTGGAGTGGAGT 62311  
 1315 AATCGGATCAGATGCTGCGGTAATGATGCTGCGGATCTTGTATACACCGGCTTCA 1374  
 62312 AATCGGATCAGATGCTGCGGTAATGATGCTGCGGATCTTGTATACACCGGCTTCA 62371  
 1375 CACCATGAGTGGGTTTACAGAGAGGAGTGTCTTACCGTGAAGAGGAGGAGGAGGAG 1434  
 62372 CACCATGAGTGGGTTTACAGAGAGGAGTGTCTTACCGTGAAGAGGAGGAGGAGGAG 62431  
 1435 CGGTGAATTCATGATGCTGGGTG 1457  
 62432 CGGTGAATTCATGATGCTGGGTG 62454

RESULT 45  
 AAA81490\_02  
 Continuation (3 of 15) of AAA81490 from base 200001 (N. meningitidis B full length genom  
 WP Sequence split into 15 fragments LOCUS AAA81490 Accession Aaa81490  
 WP Fragment Name Begin End  
 WP AAA81490\_00 1 110000  
 WP AAA81490\_01 100001 210000  
 WP AAA81490\_02 200001 310000  
 WP AAA81490\_03 300001 410000  
 WP AAA81490\_04 400001 510000  
 WP AAA81490\_05 500001 610000  
 WP AAA81490\_06 600001 710000  
 WP AAA81490\_07 700001 810000  
 WP AAA81490\_08 800001 910000  
 WP AAA81490\_09 900001 1010000  
 WP AAA81490\_10 1000001 1110000  
 WP AAA81490\_11 1100001 1210000  
 WP AAA81490\_12 1200001 1310000  
 WP AAA81490\_13 1300001 1410000  
 WP AAA81490\_14 1400001 1437668

Query Match 77.2%; Score 1124.6; DB 3; Length 110000;  
 Best Local Similarity 87.4%; Pred. No. 0; Mismatches 179; Indels 6; Gaps 4;  
 Matches 1278; Conservative

1 ATTGAACGCTGGCGGATGCTTTATACATGCAAGTGCAGACGCGACAC--GGATGCTTGC 58  
 107742 ATTGAACGCTGGCGGATGCTTTATACATGCAAGTGCAGACGCGACACGAGAGCTTGC 107801

59 ATCT--GGTGGAGAGTGGCGGAGCGGAGTGAATGATGATGATGATGATGATGATGATG 116  
 107802 TTCTGGGAGTGGCGGAGTGGCGGAGCGGAGTGAATGATGATGATGATGATGATGATG 107861

117 GGTAAAGCATGCAAGAGATGCTTAAATACCGATATATCTTAAAGAGAGAGAGAGAGAGATC 176  
 107862 GATTAATGATGCAAGAGATGCTTAAATACCGATATATCTTAAAGAGAGAGAGAGAGATC 107921

177 GAAAGACCTTGGGCTTTGAGAGCGGCGGATGCTGATGATGATGATGATGATGATGATG 236  
 107922 TTCTGGGAGTGGCGGAGTGGCGGAGCGGAGTGAATGATGATGATGATGATGATGATG 107981

237 TTAACCAAGGCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 296  
 107982 TTAACCAAGGCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 108041

297 CACGCGCAGAGCTCTTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 356  
 108042 CACGCGCAGAGCTCTTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 108101

357 GATTCAGCAATGCGCGGTGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 416  
 108102 GATTCAGCAATGCGCGGTGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 108161

417 AAGAAAGTTACGGTAAATATCTGATCTATGACGGTATGACAGAAAGACCGGC 476  
 108162 AAGAAAGGCTGTCTATATATCAGCGGTATGACGGTAACTGAAAGATTAAGACCGGC 108221

477 TAACTAGTGCAGACGCGCGGTAAATAGTAGGTGCAAGGGTTAAATCGGAATTACGG 536  
 108222 TAACTAGTGCAGACGCGCGGTAAATAGTAGGTGCAAGGGTTAAATCGGAATTACGG 108281

537 GCGTAAAGGAGTGCAGCGCGGTAAATAGTAGGTGCAAGGGTTAAATCGGAATTACGG 596  
 108282 GCGTAAAGGAGTGCAGCGCGGTAAATAGTAGGTGCAAGGGTTAAATCGGAATTACGG 108341

597 GAATTCGCTTTGAACTACAGGCTTAAAGTGTGCGAGAGGAGTGAATTCATGTGTA 656  
 108342 GAATTCGCTTTGAACTACAGGCTTAAAGTGTGCGAGAGGAGTGAATTCATGTGTA 108401

657 GCAATGAATTCGCTGAGATATGGAAGAACTGATGCGGAGGAGGAGTGAATTCATGTGTA 716  
 108402 GCAATGAATTCGCTGAGATATGGAAGAACTGATGCGGAGGAGGAGTGAATTCATGTGTA 108461

QY	717	CACTGACGCTCATGACGAAAGCGTGGAGCAACAGATTAGATACCTGGTAGTCCA	776
DB	108462	CACGTACGTTCAATGCCGAAAGCGTGGAGCAACAGATTAGATACCTGGTAGTCCA	108521
QY	777	CGCCCTAAACGATGTCACACTAGTGTGGGCTTTATTA-GGCTTGGTAAAGAGCTTAAG	835
DB	108522	CGCCCTAAACGATGTCACACTAGTGTGGGCAACGATTGGCTGGTAGCTAGCTTAAG	108581
QY	836	CCTGAAGTTGACCCGCTGGGAGTACGGTTCGCAAGTTAAACTCAAGAAATTGACGGG	895
DB	108582	CCTGAAGTTGACCCGCTGGGAGTACGGTTCGCAAGTTAAACTCAAGAAATTGACGGG	108641
QY	896	GACCCGCAACAAGCGGTGATTAATGTGGATTAAATTCGATCAACGCGAAAACTTACCTA	955
DB	108642	GACCCGCAACAAGCGGTGATTAATGTGGATTAAATTCGATCAACGCGAAAACTTACCTA	108701
QY	956	CCCTTGACATGTAGCGCAATTTTCTAGAGTAGATTAGTG-CTTCGGGAACGCTAACACAG	1014
DB	108702	GTCTTGACATGTAGCGCAATTTTCTAGAGTAGATTAGTG-CTTCGGGAACGCTAACACAG	108761
QY	1015	GTGCTGCATGCGCTGTCTGCACTCGTGTCTGATGATTTGGGTTAAGTCCCGCAACGAGC	1074
DB	108762	GTGCTGCATGCGCTGTCTGCACTCGTGTCTGATGATTTGGGTTAAGTCCCGCAACGAGC	108821
QY	1075	GCAACCCCTGTGATTAAATGGCATCATTTGGTTGGGCACTTTAATGAGACTGCGGGTAGC	1134
DB	108822	GCAACCCCTGTGATTAAATGGCATCATTTGGTTGGGCACTTTAATGAGACTGCGGGTAGC	108881
QY	1135	AAACCGGAGGAAGTGGGGATGACGTCAAGTCTCTCATGGCCCTTATGAGGTAGGGCTTCA	1194
DB	108882	AAACCGGAGGAAGTGGGGATGACGTCAAGTCTCTCATGGCCCTTATGAGGTAGGGCTTCA	108941
QY	1195	ACGTAAATACAAATGGCCCGCTACAGAGGGTTGCCAACCCCGGAGGGAGCTTAATCTCAAA	1254
DB	108942	ACGTAAATACAAATGGCTGCTACAGAGGGTTGCCAACCCCGGAGGGAGCTTAATCTCAAA	109001
QY	1255	ACCGGCTGTGATCCGGATCGGAGTCTGCACTCGATCGAATCGGAATCGCTAGT	1314
DB	109002	ACCGGCTGTGATCCGGATCGGAGTCTGCACTCGATCGAATCGGAATCGCTAGT	109061
QY	1315	AATCGCGATCGACATGTCGCGGTGAAATCGTTCGCCGGTCTTGTACACACCGCCGCTCA	1374
DB	109062	AATCGCGATCGACATGTCGCGGTGAAATCGTTCGCCGGTCTTGTACACACCGCCGCTCA	109121
QY	1375	CACCAATGGGAGTGGGTTTCAACAAGACAGTATGTTAAACCTTAAGAGAGCGCTTGCCA	1434
DB	109122	CACCAATGGGAGTGGGAGTATCAAGAAGTATGATTAACAACAAGAGTCCGCTTACCA	109181
QY	1435	CGGTGAGATTCACTGACTGGGGTG 1457	
DB	109182	CGGTATGCTTCACTGACTGGGGTG 109204	

RESULT 46  
AAA81490\_03  
Continuation (4 of 15) of AAA81490 from base 300001 (N. meningitidis B full length genom  
WP Sequence split into 15 fragments LOCUS AAA81490 Accession Aaa81490  
WP Fragment Name Begin End

WP	AAA81490_00	1	110000
WP	AAA81490_01	210000	210000
WP	AAA81490_02	200001	310000
WP	AAA81490_03	300001	410000
WP	AAA81490_04	400001	510000
WP	AAA81490_05	500001	610000
WP	AAA81490_06	600001	710000
WP	AAA81490_07	700001	810000
WP	AAA81490_08	800001	910000
WP	AAA81490_09	900001	1010000
WP	AAA81490_10	1000001	1110000
WP	AAA81490_11	1100001	1210000
WP	AAA81490_12	1200001	1310000
WP	AAA81490_13	1300001	1410000
WP	AAA81490_14	1400001	1437668

Query Match 77.2% Score 1124.6; DB 3; Length 110000;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 1278; Conservative 0; Mismatches 179; Indels 6; Gaps 4;

QY	1	ATTGAACGCTGCGGCATGCTTTTACATGCAAGTGCAGCGACAC--GGATGCTTGC	58
DB	7742	ATTGAACGCTGCGGCATGCTTTTACATGCAAGTGCAGCGACACGAGAACTTGC	7801
QY	59	ATCT--GGTGGAGATGGCGGAGCGGTGATGATTCGGAACCTATCCGAAGAGGG	116
DB	7802	TTCCTGGGTGGGAGATGGCGGAGCGGTGATGATTCGGAACCTATCCGAGTATGAGG	7861
QY	117	GGTAAAGATGATGATGCTAATACCGCATATACCTAAAGAGAAAGCAGGGAGTC	176
DB	7862	GATTAATGATGAAAGATCAAGTATACCGCATATACCTTGAAGAGAAAGCAGGGAGTC	7921
QY	177	GAAAGACCTTGGCTTTTGAAGCGGCGCATGCTGATTAAGCTAATGTTGGTGAAGGC	236
DB	7922	TTCGGGCTTGGCTTATTCAGCGGCGCATATTCGATTAAGTATGTTGGTGAAGGC	7981
QY	237	CTACCAAGCGCATCATGATGATGTTGTTGAGAGGACACACGCACTGGGACTGAGA	296
DB	7982	CTACCAAGCGCATCATGATGATGTTGTTGAGAGGATATCCGCCACTGGGACTGAGA	8041
QY	297	CACGGCCGAGACTCCTACGGGAGCGAGAGTGGGGAATTTTGGACAAATGGGCGAAGCT	356
DB	8042	CACGGCCGAGACTCCTACGGGAGCGAGAGTGGGGAATTTTGGACAAATGGGCGAAGCT	8101
QY	357	GATCCAGCAATGCGCGTGAAGTGAAGAGGCTTCGGTGTGAAGCTTTCACTGAG	416
DB	8102	GATCCAGCAATGCGCGTGTGAAAGAGGCTTCGGTGTGAAGAGCTTTTTCAGAGG	8161
QY	417	AAAGAAAGTTACGTTAATATCTGACTCATGACGTTATCGACAGAAAGAACCCGGC	476
DB	8162	AAAGAAAGCTGTTGCTAATATCAAGCGCTGATGACGTTAAGTGAAGTAAAGAACCCGGC	8221
QY	477	TAACTAGTGCAGAGCGCGGTAAATAGTAAAGGTGCAAGGTTAATCGAATTAATCTAG	536
DB	8222	TAACTAGTGCAGAGCGCGGTAAATAGTAAAGGTGCAAGGTTAATCGAATTAATCTAG	8281
QY	537	GCGTAAAGGTGCGAGCGGCTTGTGAAGTCAATGTGAATCCCGGGCTTAACCTGG	596
DB	8282	GCGTAAACGGGGCGAGCGGTTACTTAAGCAGATGTAATCCCGGGCTCAACCCGG	8341
QY	597	GAATTCGTTGAAATCTAACAAGCTTAAAGTGTGCGAGAGGAGTGGAAATTCATGTGT	656
DB	8342	GAATTCGTTGAAATCTGGGTGACTCGAGTGTGTCAGAGGAGGTGAATTCACAGTGT	8401
QY	657	GCAAGTAAATGCTGATGATTAAGAAACATTCGATGGCGAAGGCAAGCTCCTGGGTTAA	716
DB	8402	GCAAGTAAATGCTGATGATTAAGAAATCCGATGGCGAAGGCAAGCTCCTGGGACAA	8461
QY	717	CACGTACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGGTAGTCCA	776
DB	8462	CACGTACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGGTAGTCCA	8521
QY	777	CGCCCTAAACGATGTCACACTAGTGTGGGCTTTATTA-GGCTTGGTAAAGAGCTTAAG	835
DB	8522	CGCCCTAAACGATGTCACACTAGTGTGGGCAACGATTGGCTGGTAGCTAGCTTAAG	8581
QY	836	CCTGAAGTTGACCCGCTGGGAGTACGGTTCGCAAGTTAAACTCAAGAAATTGACGGG	895
DB	8582	CCTGAAGTTGACCCGCTGGGAGTACGGTTCGCAAGTTAAACTCAAGAAATTGACGGG	8641
QY	896	GACCCGCAACAAGCGGTGATTAATGTGGATTAAATTCGATCAACGCGAAAACTTACCTA	955
DB	8642	GACCCGCAACAAGCGGTGATTAATGTGGATTAAATTCGATCAACGCGAAAACTTACCTA	8701
QY	956	CCCTTGACATGTAGCGCAATTTTCTAGAGTAGATTAGTG-CTTCGGGAACGCTAACACAG	1014
DB	8702	GTCTTGACATGTAGCGCAATTTTCTAGAGTAGATTAGTG-CTTCGGGAACGCTAACACAG	8761

QY	1015	TTGCTGCATGCGCTGCGCAGCTCGTGTGTCGTAATGTTGGGTTAAAGTCCGCAACGAGC	1074
Db	8762	GTGCTGCATGCGCTGTCGACGTGCTGTGTCGTAATGTTGGGTTAAAGTCCGCAACGAGC	8821
QY	1075	GCAACCCCTTGTTCATTAATTGGCCATCTTGGTGTGGGCACTTAAATGAGACTGCCGTGAC	1134
Db	8822	GCAACCCCTTGTTCATTAATTGGCCATCTTGGTGTGGGCACTTAAATGAGACTGCCGTGAC	8881
QY	1135	AAACCGGAGGAAGGTGGGGATGACGTCAAGTCTTCATATGGCCCTTAATGGGTATGGGCTTCAC	1194
Db	8882	AAGCCGGAGGAAGGTGGGGATGACGTCAAGTCTTCATATGGCCCTTAATGACAGGGGCTTCAC	8941
QY	1195	ACGTATATCAATAGCGCGCTACAGAGGGTTGGCAACCCCGAGAGGGGAGACTAATCTCAGA	1254
Db	8942	ACGTATATCAATAGCTGCGGTACAGAGGGTAAGCCAGCCCGAGAGGGGAGCAATCTCAGA	9001
QY	1255	AGCGCGTGTAGTCCCGATCCGAGTCTGCAACTCGACTCCGTGAAGTGGGAATGCTGTAGT	1314
Db	9002	AACCCATGTGTAGTCCCGATCTGCAACTCTCACAACCTCAGTGCATGAAGTGGGAATGCTGTAGT	9061
QY	1315	AATGCGGATCAGCATGTGCGCGGTAAATACGTTCCGGGACTCTTGACACACCGCCCGTCA	1374
Db	9062	AATGCGAGTCAAGCATCTGCGGTAAATACGTTCCCGGACTCTTGACACACCGCCCGTCA	9122
QY	1375	CACCATGGGAGTGGGTTTCACCAAGACAGTATGCTTAAACGTAAGAGGGCGCTTGCCA	1433
Db	9122	CACCATGGGAGTGGGGSATATACAGAGTAAGTAGATATTAACCAAGAGAGTCCGCTTAACA	9181
QY	1435	CGGTGAGATTCATGACTGGGGTG	1457
Db	9182	CGGTATGCTTCATGACTGGGGTG	9204

RESULT 47  
AAA81489 2/c  
Continuation (3 of 9) of AAA81489 from base 200001 (N. meningitidis partial DNA sequences  
WP Sequence split into 9 fragments LOCUS AAA81489 Accession Aaa81489

WP	Fragment Name	Begin	End
WP	AAA81489_0	1	110000
WP	AAA81489_1	100001	210000
WP	AAA81489_2	200001	310000
WP	AAA81489_3	300001	410000
WP	AAA81489_4	400001	510000
WP	AAA81489_5	500001	610000
WP	AAA81489_6	600001	710000
WP	AAA81489_7	700001	810000
WP	AAA81489_8	800001	837096

  

Query Match	77.2%	Score 1124.6;	DB 3;	Length 110000;
Best Local Similarity	87.4%;	Pred. No. 0;		
Matches 1278;	Conservative	0;	Mismatches 179;	Indels 6; Gaps 4
Qy	1 ATTGAACGCTGGCGGCATGCTTTACATCATGCAAGTGGAAACGGCAGAC--GGATGCTTGC	58		
Db	95282 ATTGAACGCTGGCGGCATGCTTTACATCATGCAAGTGGACGGCAGACAGAAAGCTTGC	9522		
Qy	59 ATCT--GGTGGCGAGTGGCGGACGGGTGATATGATCATGCGAAAGTATCCAGAAAGCGG	116		
Db	95222 TTCTCGGATGGCGAGTGGCGAACGGGTGATTAATCATATGCGAAAGTATCCAGAAAGTGGG	9512		
Qy	117 GGTAAAGCATCGAAGATGTCTTAATACCGCATATATCTTAAGAGGAAAGCAGGGGATTC	176		
Db	95162 GATTAATCGAATAAGATCAGCTAATAACCGCATATACGCTTGGAGAGGAAAGCAGGGGACC	9512		
Qy	177 GAAAGACTTGGCGCTTTTGGAGGGCGCGATGTCGATTAGCTATGTTGGTGGGCTAAAGG	236		
Db	95102 TTGCGGCTTGGCGCTTATTCAGCGCGCGAATATCTGATTAGCTATGTTGGTGGGCTAAAGG	9502		
Qy	237 CTACCAAGCGCAGATCAGTATGTTGGTCTTGAGAGAGCAGCAACGCCACACTGGACTGAGA	296		
Db	95042 CTACCAAGCGCAGATCAGTATGTTGGTCTTGAGAGAGTATCCGCCACACTGGACTGAGA	949		
Qy	297 CACGCCCCAGATCTCTACGGGAGCAGCAGTGGGGAATTTTGGACATATGGCGCAAGCT	356		

Db	94982	CACGGCCACAGACTCTTAACGGAGGCGACAGTGGGAAATTTTGGACAAATGGGCGCAGGCT	94923
Oy	357	GATCCAGCAATGCCCGGTAGTAAGAAAGCCCTTCGGGTGTGAAGCTTTCACTGAG	9416
Db	94922	GATTCAGCCATGCGCGGTCTGTGAAGAAAGCCTTCGGGTGTGAAGCACTTTGTCAAGG	94863
Oy	417	AAGAAAAGTTTACGTAAATTAATCGTGACTCTGACGTTATGACAGAAAGAAAGCACCGG	9476
Db	94862	AAGAAAAGCGCTGTGTCTAAATATCAGCGCGCTGATGACGGTAACTGAAGATTAAGCACCGG	94803
Oy	477	TAACTACGTGCCAGCAGCCCGGTAAATACGTAGGGGTGAAGCGTTAATCCGAATTACTGG	94743
Db	94802	TAACTACGTGCCAGCAGCCCGGTAAATACGTAGGGGTGAAGCGTTAATCCGAATTACTGG	94743
Oy	537	GCGTAAAGGGTGGCAGACGCGCTTTGTGAAGTGAATGTGAAATCCCGGGCTTAACTGG	94683
Db	94742	GCGTAAAGCGGGCGCAGACCGGTACTTAAAGCAGAGTGAATATCCCGGGCTCAACCGG	94623
Oy	597	GAATTGGCTTTGAACCTACAAAGGCTTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTGA	94657
Db	94682	GAATCGCGTTCTGAACCTGGGTGACTCGAGTGTGTGAGAGGAGGTGAATTCACGTGTGA	94623
Oy	657	GCACTGAATGCGTAGAGATTAATGGAAACATCGAATGGGGAAGGCGCTCTGGGTTAA	94563
Db	94622	GCACTGAATGCGTAGAGATTAATGGAAACATCGAATGGGGAAGGCGCTCTGGGGAACA	94563
Oy	717	CACGTAGCGCTCATGCAGAAAGCGTGGGAGCAAAACAGATTAAGATACCTGTGATGCA	94503
Db	94562	CACGTAGCGCTCATGCAGAAAGCGTGGGAGCAAAACAGATTAAGATACCTGTGATGCA	94503
Oy	777	CGCCCTAAACGATGTCAACTAGTTGTTGGGCTTTATTA-GGCTTGTGAACGAAGCTAACG	94443
Db	94502	CGCCCTAAACGATGTCAACTAGTTGTTGGGCAACCTGAATGCTGTGATCGTAGCTTAAG	94443
Oy	836	CGTGAAGTTGACCGCTCTGGGGAATTAAGTTCGCAAGATTAATACTCAAAGAAATTGACGG	94389
Db	94442	CGTGAAGTTGACCGCTCTGGGGAATTAAGTTCGCAAGATTAATACTCAAAGAAATTGACGG	94389
Oy	896	GACCCGACAAAGCGGTGGAATTATGTGGATTAATTCGATCAACGCGAAAAACCTTAACCTA	94323
Db	94382	GACCCGACAAAGCGGTGGAATTATGTGGATTAATTCGATCAACGCGAAAGACCTTAACCTG	94323
Oy	956	CCCTTGAATGTAGCGAATTTTCTAGATTAAGTAAATGTC-CTTCGGGAACGCTAACACAG	94263
Db	94322	GCTTTGAATGTAGCGAATCTCTCGGAGACGAGAGAGTCCCTTCGGGAACGCTTAACACAG	94263
Oy	1015	GTGCTGCATGGCTGTGCTCAGCTCTGTGTGCTGAGATGTTGGGTAAATGCCCAACGAGC	94203
Db	94262	GTGCTGCATGGCTGTGCTCAGCTCTGTGTGCTGAGATGTTGGGTAAATGCCCAACGAGC	94203
Oy	1075	GCAACCCCTTGCAATTAATGTCATCAATTTGGTGGGCACTTAAATAGACTGCCGCTGAC	94143
Db	94202	GCAACCCCTTGCAATTAATGTCATCAATTTGGTGGGCACTTAAATAGACTGCCGCTGAC	94143
Oy	1135	AAACCGAGAGAGGTGGGGATGACGTACAGTCTCATAGGCCCTTAATGGGTAGGCTTAC	94103
Db	94142	AAACCGAGAGAGGTGGGGATGACGTACAGTCTCATAGGCCCTTAATAGCACAGGGCTTAC	94103
Oy	1195	ACGTAAATCAATGGCGGTACACAGGGTTGCCAACCCGCGAGGGGGAAGCTAAATCTAGAA	94082
Db	94082	ACGTAAATCAATGGCGGTACACAGGGTTGCCAACCCGCGAGGGGGAAGCTAAATCTAGAA	94082
Oy	1255	AGCGGTGTGATGCCGATGCGGAATCTGCAACTGCACTCCGTGAAGTCGGAATCGCTAGT	93962
Db	94022	AACCGATGTATGCCGATGCGGAATCTGCAACTCTGCAACTGCAATGCGGAATCGCTAGT	93962
Oy	1315	AATCCGAGATAGAGATGTCCGAGTAATAGCTTCCCGAGTCTTGTACACACCGCCGTCA	93902
Db	93962	AATCCGAGATAGAGATGTCCGAGTAATAGCTTCCCGAGTCTTGTACACACCGCCGTCA	93902
Oy	1375	CACCATGGAGTGGGTTTCAACCAAGACAGTATGCTTAACCTGAAGAGGGCGCTTGGCA	1434







QY	1	ATTGAAAGCTGTGGCGGCAATGCTTTTACAATGCAATGCAAGTGGAAAGGGAGAAC--GGATGCTTGC	58
Db	2153	ATTGAAAGCTGTGGCGGCAATGCTTTTACAATGCAATGCAAGTGGAAAGGGAGAAC--GGATGCTTGC	2094
QY	59	ATCT--GGTGGCGAGTGGCGGAGCGGATGAGTAAATGCAATCGGAACGTAATCCAAAGAGGGG	116
Db	2093	TTTCTCGGGTGGCGAGTGGCGGAGCGGATGAGTAAATGCAATCGGAACGTAATCCAAAGAGGGG	2034
QY	117	GGTAAAGCATGGAAGAATGTGCTAATATACCGCATATACTCTAAGAGAGAAAGCGGGATC	176
Db	2033	GATAATCATGATGGAAGAATCAGACTAATCCCATCTCGTCTTGAGAGAGAAAGCGGGAGACC	1974
QY	177	GAAAGACTTGGCGCTTTTGGAGCGGCGGAGTCTGATTAAGTCTGTTGGTGGGGTTAAAGG	236
Db	1973	TTTCGGGCGCTTGGCGCTTATTCGAGCGGCGGAGTCTGATTAAGTCTGTTGGTGGGGTTAAAGG	1914
QY	237	CTACCAAGAGCGAGCATAGTAGTGGTCTGAGAGAGACGACAGCAGCACAATTGGGACTTGAGA	296
Db	1913	CTACCAAGAGCGAGCATAGTAGTGGGCTCTGAGAGAGATGATCCGCAACACTGGGACTTGAGA	1854
QY	297	CACGGCCCAAGCTCTTACGGGAGGACAGATGSGGAAATTTTGGACAATGGGCGCAGCCT	356
Db	1853	CACGGCCCAAGCTCTTACGGGAGGACAGATGSGGAAATTTTGGACAATGGGCGCAGCCT	1794
QY	357	GATCCAGCAATGCCCGCGTAGTGAAGAAGGCTTCGGGTTGTAAAGCTTTTCAAGTCGAG	416
Db	1793	GATCCAGCAATGCCCGCGTAGTGAAGAAGGCTTCGGGTTGTAAAGCACTTTGTGACGG	1734
QY	417	AAGAAAGAGTTACGGTAAATTAATCGTAGCTCATGACGGTATCGACAGAGAAGACCGGC	476
Db	1733	AAGAAAGAGCTGTTGCTAATATCAGCGGCTATACGGTACTGGAAGATTAAGACCGGC	1674
QY	477	TAATCATCGTGCAGAGACCGCGGTAAATCGTAGGGTSCAAGCGTTAATCGAATTACTGG	536
Db	1673	TAATCATCGTGCAGAGACCGCGGTAAATCGTAGGGTSCAGCGTTAATCGAATTACTGG	1614
QY	537	GCGTAAAGGGTGGCGAGCGCGCTTTGTAAATGCAATGTGAAATCCCGGGCTTAAACCTGG	596
Db	1613	GCGTAAACCGGGCGGAGCGGTTACTTAAAGAGAGATGTGAATCCCGGGCTCAACCGG	1554
QY	597	GAAATGCGTTTGAATCTAAGGGCTAGAGTGTGSCAGAGGAGGTGGAATTCATGTGTA	656
Db	1553	GAACTGCGTTCTGAACTGGGTGACTCGAGTGTTCAGAGGAGGTGAATTCACAGTGTA	1494
QY	657	GCAGTGAATGCGTGAAGATATGGAAGAACATCGATGCGAAGCGACTCTCTGGGTTAA	716
Db	1493	GCAGTGAATGCGTGAAGATGTGAGAGAAATCCCATGCGAAGCGACTCTCTGGGCAA	1434
QY	717	CAGTGAACGCTATGCAAGAAAGCGTGGGAGCAAAACAGATTAATACCTGTGATGCA	776
Db	1433	CAGTGAACGTTATGTCCGAAAGCGGTGTAGCAAAACAGATTAATACCTGTGATGCA	1374
QY	777	GGCCCTAAACGATGTCAAATGATGTTGGGCGCTTATTA--GGCTTGTGTAAAGAACTTAAGC	835
Db	1373	GGCCCTAAACGATGTCAAATTAATCTGTTGGGCAACTGATGCTTGTGTAGCTTAAGC	1314
QY	836	CGTGAAGTTGACCGGCTGGGGAGTACCGTCCGCAATTTAAACTCAAGAAATGACGGG	895
Db	1313	CGTGAATTTGACCGGCTGGGGAGTACCGGTCCGCAATTTAAACTCAAGAAATTTGACGGG	1254
QY	896	GACCCGCAAGAGCGGTGATTATGTGTGATTATTTGATGCAACGGGAAAACTTAACTTA	955
Db	1253	GACCCGCAAGAGCGGTGATGTGTGATTATTTGATGCAACGGGAAAGAACTTTAACTTG	1194
QY	956	CCCTTGACATGTAGGGAATTTTCTAGAGATAGATTAGTG--CTTCGGAGACGCTAACAG	1014
Db	1193	GCTTTGACATGTAGGGAATCTTCGAGAGCGAGAGATGCTTTCGGAGCCGTAAACAG	1134
QY	1015	GTCCTGCATAGGCTGTGCTCAGCTCGTGTGCTGAGATGTTGGTTAAGTCCGCAAGAC	1074
Db	1133	GTCCTGCATAGGCTGTGCTCAGGCTCGTGTGCTGAGATGTTGGGTTAAGTCCGCAAGAC	1074
QY	1075	GCAACCTTGTATTAATTGCCATCATTTGGTTGGGCACTTTAATGAACTGCGGGTAC	1134

Db	1073	GCAACCCCTTGCATTAGTTCGCATCATTCAGTTGGGACCTCTAATGAGACTGCCGGTGAC	1014
Qy	1135	AAACCGAGGAGGTGGGGATGACGTCAAGTCTCATGGCCCTTATGGTAGGGCTTCAC	1194
Db	1013	AAAGCCGAGGAGGTGGGGATGACGTCAAGTCTCATGGCCCTTATGACGAGGCTTCAC	954
Qy	1195	ACGTATATCAATGGCGCGTACAGAGGGTTGCGCAACCCGCGAGGGGAGACTATCTCAGAA	1255
Db	953	ACGTATATCAATGGTTCGGTACAGAGGGTAGCAGCGCGAGGGCGAGCCATCTCAAA	894
Qy	1255	AGCGCGTGTAGTCCGGATCGGAGTCTGCAACTGCACTCCGTGAAGTCGGAAATGGCTAGT	1314
Db	893	AAACGATGTAGTCCGGATTCGACTCTGCGCACTGAGTGCATGAAGTCGAATGCTAGT	834
Qy	1315	AATGGCGGATACGATGTGCGGGGTGAATAGCTTCCGGGCTTTGTACACACGCCCGTCA	1374
Db	833	AATGGCAGGTACGATACGTGGGGGAATACGTTCCCGGGCTTTGTACACACGCCCGTCA	774
Qy	1375	CACCATGGGAGTGGGGTTTCAACAGAGCAGGTAGCTTAACCGTAAGAGGGCGCTTGCCA	1435
Db	773	CACCATGGGAGTGGGGGATACCAAGATGATAGGATTAACCAAGAGAGTCCGCTTACCA	714
Qy	1435	CGGTGAGTTCATGACTGGGGTG	1457
Db	713	CGGTATGCTTCATGACTGGGGTG	691

RESULT 50  
AAAF21613/c  
ID AAAF21613 standard; DNA; 172325 BP.

DT 13-MAR-2001 (first entry)

DE *Neisseria meningitidis* B nucleotide sequence SEQ ID NO:114.

KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;  
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
 KW ds.

OS *Neisseria meningitidis*.

PN WO200066791-A1.

PD 09-NOV-2000.

08-MAR-2000; 2000WO-US005928.

PR 30-APR-1999; 99US-0132068P.

PR 28-FEB-2000; 2000GB-00004695.

PA (CHIR ) CHIRON CORP.

3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527

Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;

XX

[illegible]

frames are used to detect, treat and prevent Neisserial infections.

PS Claim 7; Appendix A; 692pp; English

AA The present invention describes the full length genome of *Neisseria*  
CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613  
CC represent fragments of the NMB genomic sequence, as the sequence was too  
CC long to go in a record on its own it was split into 8 sequences which  
CC overlap each other at the beginning and end of each sequence by 49980 bp



PR 30-APR-1999; 99US-0132068P.  
 PR 08-OCT-1999; 99MO-US023573.  
 PR 28-FEB-2000; 2000GB-00004695.  
 PA (CHIR ) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 PI Pizza M, Hickey E, Peterson J, Tetelín H, Venter JC;  
 PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
 PI Rappuoli R, Frazer CM, Grandi G;  
 DR WPI; 2000-647603/62.

Neisseria meningitidis B full length genome sequence and open reading  
 frames are used to detect, treat and prevent Neisserial infections.

Claim 7; Appendix A; 692pp; English.

The present invention describes the full length genome of Neisseria  
 meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613  
 represent fragments of the NMB genomic sequence, as the sequence was too  
 long to go in a record on its own it was split into 8 sequences which  
 overlap each other at the beginning and end of each sequence by 4980 bp  
 (i.e. the last 4980 bp of AAF21544 is repeated at the beginning of  
 AAF21607, the last 4980 bp of AAF21607 are repeated at the beginning of  
 AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins  
 given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR  
 primers which are used in the exemplification of the present invention.  
 The NMB genome and fragments from it have antibacterial activity, and can  
 be used in vaccines and gene therapy. Neisseria nucleic acids, proteins  
 and/or antibodies which binds to the proteins can be used in compositions  
 for treating or preventing infection due to Neisserial bacteria or as a  
 diagnostic reagent for detecting the presence of Neisserial bacteria or  
 of antibodies raised to Neisserial bacteria. Computers, computer memory,  
 computer storage medium or computer databases can be used in a search to  
 identify open reading frames (ORFs) or coding sequences within the NMB  
 genome. The DNA sequences provide further opportunities to find antigenic  
 or immunogenic proteins which are more effective in vaccines than the  
 outer membrane proteins currently used

Sequence 349980 BP; 84410 A; 84863 C; 94187 G; 86520 T; 0 U; 0 Other;

Query Match 77.2%; Score 1124.6; DB 3; Length 349980;  
 Best Local Similarity 87.4%; Pred. No. 0;  
 Matches 1278; Conservative 0; Mismatches 179; Indels 6; Gaps 4;

QY 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAAGTGCAGACGCGAC--GGATGCTTGC 58  
 DB 7742 ATTGAACGCTGGCGGCGATGCTTTACATGCAAGTGCAGACGCGACGAGAAAGCTTGC 7801  
 QY 59 ATCTT--GGTGGCGAGTGGCGGCGAGCGGAGTAATGCAATCGAAACGTATCCAGAAAGAGGG 116  
 DB 7802 TTCTCGGGTGGCGAGCGGCGAGCGGAGTAATGCAATCGAAACGTATCCAGAAAGAGGG 7861  
 QY 117 GGTAAACCATCGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAACAGAGGGATC 176  
 DB 7862 GATPACTGATCGAAAGATCACTAATACCGCATATCTTAAGAGAGAAAGAGGGAGCC 7921  
 QY 177 GAAAGACCTTGGCTTTTGAAGCGCGGAGTGTGATTTAGTTAGTTGGTGGGTTAAAGGC 236  
 DB 7922 TTGGGGCTTGGCTTATTCGAGCGCGCGCATATCTGATTAGTTAGTTGGTGGGTTAAAGGC 7981  
 QY 237 CTACCAAGGCGAGCATAGTAATGTTGTTGAGAGAGAGACGACACCTGGAGCTGAGA 296  
 DB 7982 CTACCAAGGCGAGCATAGTAATGTTGTTGAGAGAGAGATCCGACACTGAGAGCTGAGA 8041  
 QY 297 CACGGCCCAAGACTCTTACGAGGAGCAGCATGTGGGAAATTTTGAACAATGGCGCAAGCCT 356  
 DB 8042 CACGGCCCAAGACTCTTACGAGGAGCAGCATGTGGGAAATTTTGAACAATGGCGCAAGCCT 8101  
 QY 357 GATCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTTAAAGCTTTTCAGTGAAG 416  
 DB 8102 GATCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTTAAAGCACTTTTTCAGGAG 8161

QY 417 AAGAAAAGTTACGCTTAATATGTCATGACGCTATCGACAGAAAGACACCGGC 476  
 DB 8162 AAGAAAAGCTGTTGCTTAATATGTCATGACGCTATCGACAGAAAGTAAAGACACCGGC 8221  
 QY 477 TAACTACGTGCGCAGACCGCGGTAAATGTAAGGTGCAAGCGTTAATCGAATTAATCTGG 536  
 DB 8222 TAACTACGTGCGCAGACCGCGGTAAATGTAAGGTGCAAGCGTTAATCGAATTAATCTGG 8281  
 QY 537 GCGTAAAGGTGCGCAGCGCGCTTTGTAATGTCATGTAAGTAAATCCCGGGCTTAATCTGG 596  
 DB 8282 GCGTAAAGGTGCGCAGCGCGCTTTGTAATGTCATGTAAGTAAATCCCGGGCTTAATCTGG 8341  
 QY 597 GAATTCGCTTGAATCAAGCTCAAGGCTAGAGTGGCAGAGGAGTGAATTCATGTGTA 656  
 DB 8342 GAATTCGCTTGAATCAAGCTCAAGGCTAGAGTGGCAGAGGAGTGAATTCATGTGTA 8401  
 QY 657 GCAGTGAATGCGTGAAGATATGAAAGAACATCATGATGGGAAAGGACGCTCTGGGTTAA 716  
 DB 8402 GCAGTGAATGCGTGAAGATATGAAAGAACATCATGATGGGAAAGGACGCTCTGGGTTAA 8461  
 QY 717 CACTGACGCTATGACAGAAACGCTGGGAGCAAAAGATTAATGATACCTGTGATGCA 776  
 DB 8462 CACTGACGCTATGACAGAAACGCTGGGAGCAAAAGATTAATGATACCTGTGATGCA 8521  
 QY 777 CGCCTTAAACGATGTCATCTACTTGTGGGCTTATTA--GGCTGGTAAAGAACTTAACG 835  
 DB 8522 CGCCTTAAACGATGTCATCTACTTGTGGGCTTATTA--GGCTGGTAAAGAACTTAACG 8581  
 QY 836 CGTGAAGTTGACCGCGCTGGGAGTACCGTTCGCAAGATTAAACCTCAAGAAATTGACGG 895  
 DB 8582 CGTGAAGTTGACCGCGCTGGGAGTACCGTTCGCAAGATTAAACCTCAAGAAATTGACGG 8641  
 QY 896 GACCCGCAACGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 955  
 DB 8642 GACCCGCAACGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 8701  
 QY 956 CCGTTGACATGTAAGCAATTTTCTAGATTAATGATTAATGATTAATGATTAATGATTA 1014  
 DB 8702 GCTTGAACATGTAAGCAATTTTCTAGATTAATGATTAATGATTAATGATTAATGATTA 8761  
 QY 1015 GGTGTCATGAGGCTGTCAGTCTGTCAGTCTGTCAGTCTGTCAGTCTGTCAGTCTGTCAG 1074  
 DB 8762 GGTGTCATGAGGCTGTCAGTCTGTCAGTCTGTCAGTCTGTCAGTCTGTCAGTCTGTCAG 8821  
 QY 1075 GCAACCTTGTCAATTAATGTCATCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1134  
 DB 8822 GCAACCTTGTCAATTAATGTCATCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 8881  
 QY 1135 AAACCGAGAGAGGTGGGATGACGTCAGTCTTATGATGATGATGATGATGATGATGATG 1194  
 DB 8882 AAACCGAGAGAGGTGGGATGACGTCAGTCTTATGATGATGATGATGATGATGATGATG 8941  
 QY 1195 AACTTAATCAATGCGCGTACAGAGGTTGCCAACCGCGAGGGAGGAGCTTAATCTGAGA 1254  
 DB 8942 AACTTAATCAATGCGCGTACAGAGGTTGCCAACCGCGAGGGAGGAGCTTAATCTGAGA 9001  
 QY 1255 AGCGGCTGTAATGTCAGGATGAGTCTGCAACTGCACTCGTGAAGTGGAAATGCTGAGT 1314  
 DB 9002 AACTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9061  
 QY 1315 AATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1374  
 DB 9062 AATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9121  
 QY 1375 CACATGCGAGTGGGTTTCCAGAAAGCAGTGTGTAACCTGAAGAGAGGCGCTTGGCA 1434  
 DB 9122 CACATGCGAGTGGGTTTCCAGAAAGCAGTGTGTAACCTGAAGAGAGGCGCTTGGCA 9181  
 QY 1435 CCGTGAAGTTCATGACTGGGGTG 1457  
 DB 9182 CCGTGAAGTTCATGACTGGGGTG 9204

RESULT 52  
AAFP21611/C  
ID AAFP21611 standard; DNA; 349980 BP.  
XX  
AC AAFP21611;  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;  
diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
de.  
XX  
OS Neisseria meningitidis.  
XX  
PN WO200066791-A1.  
XX  
PD 09-NOV-2000.  
XX  
PF 08-MAR-2000; 2000WO-US005928.  
XX  
PR 30-APR-1999; 99US-0132068P.  
PR 08-OCT-1999; 99WO-US023573.  
PR 28-FEB-2000; 2000GB-00004695.  
XX  
PA (CHIR) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Pizze M, Hickey B, Peterson J, Tetrelin H, Venter JC;  
Maggiari V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
PI Rapuoli R, Frazer CM, Grandi G;  
XX  
DR WPI; 2000-647603/62.  
XX  
PT Neisseria meningitidis B full length genome sequence and open reading  
frames are used to detect, treat and prevent Neisserial infections.  
XX  
PS Claim 7, Appendix A; 692pp; English.  
XX  
CC The present invention describes the full length genome of Neisseria  
meningitidis B (NMB). The sequences in AAFP21544 and AAFP21607 to AAFP21613  
represent fragments of the NMB genomic sequence, as the sequence was too  
long to go in a record on its own it was split into 8 sequences which too  
overlap each other at the beginning and end of each sequence by 49980 bp  
(i.e. the last 49980 bp of AAFP21544 is repeated at the beginning of  
AAFP21607, the last 49980 bp of AAFP21607 are repeated at the beginning of  
AAFP21608, and so on). AAFP21545 to AAFP21588 encode the Neisseria proteins  
given in AAB58550 to AAB58593, and AAFP21589 to AAFP21606 represent PCR  
primers which are used in the exemplification of the present invention.  
CC The NMB genome and fragments from it have antibacterial activity, and can  
be used in vaccines and gene therapy. Neisseria nucleic acids, proteins  
and/or antibodies which binds to the proteins can be used in compositions  
for treating or preventing infection due to Neisserial bacteria or as a  
diagnostic reagent for detecting the presence of Neisserial bacteria or  
of antibodies raised to Neisserial bacteria. Computers, computer memory,  
CC computer storage medium or computer databases can be used in a search to  
identify open reading frames (ORFs) or coding sequences within the NMB  
genome. The DNA sequences provide further opportunities to find antigenic  
CC or immunogenic proteins which are more effective in vaccines than the  
outer membrane proteins currently used  
XX  
SQ Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 U; 0 Other;

Query Match 77.2%; Score 1124.6; DB 3; Length 349980;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 1278; Conservative 0; Mismatches 179; Indels 6; Gaps 4;

QY 1 ATTGAACGCTGGCGGCGATCTTTACACATGCAAGTGAACGGCAGAC--GGATGCTTGC 58  
DB 230511 ATTGAACGCTGGCGGCGATCTTTACACATGCAAGTGAACGGCAGACGAGAAGCTTGC 230452

QY 59 ATCT--GGTGGCGAGTGGCGGACGGGTGATGATATGATCGAATCGAATCGAATCGAAGAGGGG 116  
DB 230451 TTCTCGGGTGGCGAGTGGCGAACGGGTGATGATATCGAATCGAATCGAATCGAAGAGGGG 230392  
QY 117 GGTAAAGCATGGAAGATGCTCTAATACCGCATATCTCTAAGAGGAAGAGAGGGGATC 176  
DB 230391 GATAACTGATGGAAGATGATCATATATCCGATATGCTTGAAGAGGAAGAGAGGGGAGC 230332  
QY 177 GAAAGACCTTGGCGCTTTTGGAGCGGCGGATGCTGATTTAGCTAGTTGGTGGGTTAAAGC 236  
DB 230331 TTCCGGGCTTGGGCTATTTCGAGCGGCGCATATCTGATTAGCTAGTTGGTGGGTTAAAGC 230272  
QY 237 CTACCAAGCCGACGATCATGATGTTGCTGAGAGACGACCAAGCCACATGGGACTGAGA 296  
DB 230271 CTACCAAGCCGACGATCATGATGTTGCTGAGAGATGATCCGACACATGGGACTGAGA 230212  
QY 297 CACGGCCGACACTCTCTACGAGAGGACGAGAGTGGGGAATTTTGGACATGGGCGCAAGCT 356  
DB 230211 CACGGCCGACACTCTCTACGAGAGGACGAGAGTGGGGAATTTTGGACATGGGCGCAAGCT 230152  
QY 357 GATCCAGCAATGCCCGGATGGAAGAGAGGCTTCCGGTGTAAAGCTTTCAAGTCGAG 416  
DB 230151 GATCCAGCAATGCCCGGATGGAAGAGAGGCTTCCGGTGTAAAGCTTTTGTCAAGG 230092  
QY 417 AAGAAAGGTTACGGTAATATATCGTGAATCATGACGATATCGACAGAAAGACCCGAC 476  
DB 230091 AAGAAAGGCTGTTGCTAATATATCAGCGGCTATGACGTTACTGAAGATTAAGACCCGG 230032  
QY 477 TAACTACGTCGACAGACCGCGGTATATCGTAAAGGTGCAACGCTTAATCGAATTAATCTG 536  
DB 230031 TAACTACGTCGACAGACCGCGGTATATCGTAAAGGTGCAACGCTTAATCGAATTAATCTG 229972  
QY 537 GCGTAAAGGTCGCGAGCGGCTTTGTAAGTCAGATGGAATCCCGGGCTTAACCTGG 596  
DB 229971 GCGTAAAGCGGAGCGAGCGGCTTTGTAAGTCAGATGGAATCCCGGGCTTAACCTGG 229912  
QY 597 GAATTCGCTTTGAAACTCAAGGCTAGAGTGGCAGAGGAGGTGAATTCATGTGTA 656  
DB 229911 GAATTCGCTTTGAAACTGGGTGATCTCAGTGTGTGACAGGAGGTGAATTCATGTGTA 229852  
QY 657 GCAGTGAATTCGTAAGATATGGAAGACATCGATGCGAAGGACGCTCTCTGGGTTAA 716  
DB 229851 GCAGTGAATTCGTAAGATATGGAAGACATCGATGCGAAGGACGCTCTCTGGGTTAA 229792  
QY 717 CACTGACCTCATGACAGAAAGCGTGGGAGCAACAGATTAATGATACCTGGTATGCA 776  
DB 229791 CACTGACCTCATGACAGAAAGCGTGGGAGCAACAGATTAATGATACCTGGTATGCA 229732  
QY 777 CGCCCTAAGCATGCACTAGTTGTTGGGCTTATTA--GGCTTGTAAACGAAGCTAACG 835  
DB 229731 CGCCCTAAGCATGCACTAGTTGTTGGGCTTATTA--GGCTTGTAAACGAAGCTAACG 229672  
QY 836 CGTGAAGTTGACCGCTTGGGAGTACGCTGCGAAGATTAATACTCAAGAAATTGACGGG 895  
DB 229671 CGTGAAGTTGACCGCTTGGGAGTACGCTGCGAAGATTAATACTCAAGAAATTGACGGG 229612  
QY 896 GACCCGCAACAGCGGTGATATATGGAATTAATTCGATGCAACGCGAAGAAACCTTACCTA 955  
DB 229611 GACCCGCAACAGCGGTGATATATGGAATTAATTCGATGCAACGCGAAGAAACCTTACCTG 229552  
QY 956 CCCTTGACATGAGGAATTTTCTAGAGATGATTAAGT--CTTGGGGAACGCTAACACAG 1014  
DB 229551 GCTTTGACATGAGGAATTTTCTAGAGATGATTAAGT--CTTGGGGAACGCTAACACAG 229492  
QY 1015 GTGCTGATGAGCTGTGCTGACCTGCTGCTGAGATGTTGGGTTAAGTCCGCAACGAGC 1074  
DB 229491 GTGCTGATGAGCTGTGCTGACCTGCTGCTGAGATGTTGGGTTAAGTCCGCAACGAGC 229432  
QY 1075 GCAACCTTGTGATTAATTTGGATCATATTTGGTGGGCACTTTAATGACATGCGGCTAC 1134  
DB 229431 GCAACCTTGTGATTAATTTGGATCATATTTGGTGGGCACTTTAATGACATGCGGCTAC 229372  
QY 1195 AAACCGAGGAAGGTGGGATGATGATCAAGTCTTCATGAGCCTTAATGGGTAGGCTTAC 1194



Db	336602	CGCCCTAAACGATGTCATTTACTGTTGGGCAACCTGATTTGCTTGGTAGCTGAACG	336553
Qy	836	CGTGAAGTTGACCGCGCTGGGGAGTACGTCGCAAGATTAAACTCAAGGAATTGACGG	895
Db	336542	CGTGAATTTGACCGCGCTGGGGAGTACGGTCCGAAAGATTAAACTCAAGGAATTGACGG	336483
Qy	886	GACCCGACACAGCCGGTGGATTATATGTCGATTTAATTTCGATTCGAACGCGAAAACTTACCTA	955
Db	336482	GACCCGACACAGCCGGTGGATTATGTCGATTTAATTTCGATTCGAACGCGAAAACTTACCTG	336423
Qy	956	CCCTTGAACATGTAAGCGAATTTTCTAGAGATGATTTAGTG-CTTCGGGAACGCTTAACAG	1014
Db	336422	GCTTTGACATGTACGGAATCTCTCCGAAACGAGAGAGTGCCTTCGGGAGCCGTAACACAG	336353
Qy	1015	GTCGTGACATGCGCTGTCAGCTCGTGTGTCGAGATGTTGGTTAAGTCCCGCAAGAC	1074
Db	336352	GTCGTGACATGCGCTGTCAGCTCGTGTGTCGAGATGTTGGTTAAGTCCCGCAAGAC	336303
Qy	1075	GCAACCCCTTGTGATTAATTCGATCACTTTGTTGGGCACTTTATAGAGCTGCGGTGAC	1134
Db	336302	GCAACCCCTTGTGATTAATTCGATCACTTTGTTGGGCACTTTATAGAGCTGCGGTGAC	336243
Qy	1135	AAACCGGAGGAGGTCGGGATACAGCGTCAAGCTCTCAATGCGCCCTTATGGGAGGCTTCA	1194
Db	336242	AAACCGGAGGAGGTCGGGATACAGCGTCAAGCTCTCAATGCGCCCTTATGACAGGAGCTTCA	336183
Qy	1195	ACGTATATCAATGTCGCGCGCTTACAGAGGGTTCGCAACCCCGAGGGGAGCTAATCTCAGA	1254
Db	336182	ACGTATATCAATGTCGCGCGCTTACAGAGGGTTCGCAACCCCGAGGGGAGCTAATCTCACA	336123
Qy	1255	AGCGCGTCGTAGTCCGGATCGGAGTTCGAACTCGACTCCGTGAAGTCGGAATCGCTAGT	1314
Db	336122	AACCATCGTATCGGATTCGAACTCTCGAACTCGAGTCGATGAACTCGGAATCGCTAGT	336063
Qy	1315	AATCGCGATCAAGCATGTCGCGGTGAATACGTTCCCGGGCTTGTACACACCGCCCGTCA	1374
Db	336062	AATCGCGATCAAGCATGTCGCGGTGAATACGTTCCCGGGCTTGTACACACCGCCCGTCA	336003
Qy	1375	CACCATGGAGTCGGTTTTCACCAAGACAGTAGTCTAAACCGTAAAGAGGCGCTTGGCA	1434
Db	336002	CACCATGGAGTCGGGCGGATACCAAGAGTAGTAGTAACCAAGAGAGTCGCTTACCA	335943
Qy	1435	CGGTAGATTCAATGACTGGGGTG 1457	
Db	335942	CGGTATGCTTCATGACTGGGGTG 335920	
RESULT 54			
AAAF21544			
XX	ID	AAAF21544	standard; DNA; 349980 BP.
XX	AC	AAAF21544;	
XX	DT	13-MAR-2001	(first entry)
XX	DE	Neisseria meningitidis B nucleotide sequence SEQ ID NO:1.	
XX	KM	Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;	
XX	KW	diagnosis; antigen; detection; infection; gene therapy; antibacterial;	
XX	OS	Neisseria meningitidis.	
XX	PN	WO20006791-A1.	
XX	PD	09-NOV-2000.	
XX	PF	08-MAR-2000; 2000WO-US005928.	
XX	PR	30-APR-1999; 99US-0132068P.	
XX	PR	08-OCT-1999; 99WO-US023573.	
XX	PR	28-FEB-2000; 2000GB-00004695.	
XX			

PA (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX Pizza M, Hickey E, Peterson J, Tettelein H, Venter JC,  
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
PI Rappunoli R, Frazer CM, Grandi G;  
XX WPI; 2000-647603/62.  
DR  
XX  
XX Neisseria meningitidis B full length genome sequence and open reading  
PT frames are used to detect, treat and prevent Neisserial infections.  
XX  
PS Claim 7; Appendix A; 692pp; English.

The present invention describes the full length genome of *Neisseria meningitidis* B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613 represent fragments of the NMB genomic sequence, as the sequence was too long to go in a record on its own it was split into 8 sequences by 49980 bp overlap each other at the beginning and end of each sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the *Neisseria* proteins given in AAB5550 to AAB58593, and AAF21589 to AAF21606 represent PCR primers which are used in the exemplification of the present invention. The NMB genome and fragments from it have antibacterial activity, and can be used in vaccines and gene therapy. *Neisseria* nucleic acids, proteins and/or antibodies which binds to the proteins can be used in compositions for treating or preventing infection due to *Neisseria* bacteria or as a diagnostic reagent for detecting the presence of *Neisseria* bacteria or of antibodies raised to *Neisseria* bacteria. Computers, computer memory, computer storage medium or computer databases can be used in a search to identify open reading frames (ORFs) or coding sequences within the NMB genome. The DNA sequences provide further opportunities to find antigenic or immunogenic proteins which are more effective in vaccines than the outer membrane proteins currently used

Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 U; 0 Other;

Query Match	77.28;	Score 1124.6;	DB 3;	Length 349980;
-------------	--------	---------------	-------	----------------

Matches 1278; Conservative 0; Mismatches 179; Indels 6; Gaps 4;

1 ATTGAACGCTGGCGCATGCTTTACACATGCAAGTCGAACGGCAGCAC--GGATGCTTGC 58

Db 307742 ATTGACGCTGGCGCATGCTTACACATGCAAGTCGGACGGCAGCACAGAGAAGCTTGC 307801

59 ATCT--GGTGGCAGTGGCGGACGGGTGACTAATGCATCGAACGTATCCAGAAGAGGGC 116

Db 307802 TTCTCGGGTGGCGAGTGGCGAACGGGTGAGTACATATCGGAACGTACCGAGTAGTGGG 307861

117 GGTACCGCATCGAAGATGTGCTAATACCGCATATACTCTAAGGAGGAAAGCAGGGATC 176

Db 307862 GATACTGATCGAAGATCAGCTAATACCGCATACGCTTGTGAGAGAGAAAGCAGGGACC 307921

177 GAAGACCTTGGCTTTTGAGCGCGCGATGTCGATTAGCTAGTTGCTGGGTAAGGC 236

Db 307922 TTCGGCCTTGGCTATTCGAGCGGCCGATATCTGATTAGCTAGTTGCTGGGTAAGGC 307981

237 CTACCAAGCGACGATCAGTAGTTGGTCTGAGAGGACGACCACTGGGACTGAGA 296

Db 307982 CTACCAAGCGACGATCAGTAGCGGCTCTGAGAGGATGATCCGCCACACTGGGACTGAGA 308041

297 CACGGCCAGACTTCTACGGAGGCAGCAGTGGGAATTGGACAATGGCGCAAGCCT 356

Db 308042 CACGCCCACTCTACGGAGGCACGACTGGGAATTTGGACAATGGCGCAAGCCT 308101

357 GATCCAGCAATGCCGCGTGAGTGAAGAAGGCCCTTCGGGTGTAAAGCTCTTCAGTCGAG 416

Db 308102 GATCCAGCCATGCCGCTGCTGAAGAAGGCCCTTCGGGTTGTAAAGACTTTTGTCAAGG 308161

417 AAGAAAGGTTACGGTAATAATCGTACTCATGACGGTATCGACAGAGAAGCACC GC 476

Db 308162 AAGAAAGGCTGTTGCTAATATCAGCGGCTGATGACGCTACCTGAAGATAAGCACCGGC 308221

QY 477 TAACTACGTGCCAGCAGCCGCGGTAAATAGTAAAGGTGCAAGCGTTAATCGAATTAATCTGG 536  
 DB 308222 TAACTACGTGCCAGCAGCCGCGGTAAATAGTAAAGGTGCAAGCGTTAATCGAATTAATCTGG 308281  
 QY 537 GCGTAAAGGGTGGCGAGCGGCTTTGTAAAGTCAAGTGTGAATATCCCGGGCTTAATCTGG 596  
 DB 308282 GCGTAAAGCGGGCGAGCGGCTTTGTAAAGTCAAGTGTGAATATCCCGGGCTTAATCTGG 308341  
 QY 597 GAATGGCGTTGAACTCAAGGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTGA 656  
 DB 308342 GAATGGCGTTGAACTGGGTACTCGAGTGTGTCAAGAGGAGGTGAATTCATGTGTGA 308401  
 QY 657 GCAGTGAATGCGTGAAGATATATGAAGAACATCGATGGGAAAGCAGCTCTGGGTGA 716  
 DB 308402 GCAGTGAATGCGTGAAGATATATGAAGAACATCGATGGGAAAGCAGCTCTGGGTGA 308461  
 QY 717 CACTGACGCTCATGCAAGAAAGCTGGGAGCAAAACAGATTAAGATCCCTGGTATGCA 776  
 DB 308462 CACTGACGCTCATGCAAGAAAGCTGGGAGCAAAACAGATTAAGATCCCTGGTATGCA 308521  
 QY 777 GCGCCCTAAACGATGCACTAGTGTGGGCTTATTA-GGCTTGTAAAGAAAGTGAACG 835  
 DB 308522 GCGCCCTAAACGATGCACTAGTGTGGGCTTATTA-GGCTTGTAAAGAAAGTGAACG 308581  
 QY 836 CGTGAAGTTGACCGCTGGGAGTACGTCGCAAGATTAATACTCAAGAAATGAACGG 895  
 DB 308582 CGTGAAGTTGACCGCTGGGAGTACGTCGCAAGATTAATACTCAAGAAATGAACGG 308641  
 QY 896 GACCCGCAAGAGCGGTGATTAATGTGGATTAATTCATGCAAGCGCAAAACCTTACCTA 955  
 DB 308642 GACCCGCAAGAGCGGTGATTAATGTGGATTAATTCATGCAAGCGCAAAACCTTACCTA 308701  
 QY 956 CCCTTGAACATGTAGCGCAATTTCTAGAGATTAATGTCGCGGAAAGCTTAACACAG 1014  
 DB 308702 CCCTTGAACATGTAGCGCAATTTCTAGAGATTAATGTCGCGGAAAGCTTAACACAG 308761  
 QY 1015 GTGCTCATGAGCTGTGTCAGCTCGTGTGTCAGATGTTGGTTAAAGTCCCGCAACGAC 1074  
 DB 308762 GTGCTCATGAGCTGTGTCAGCTCGTGTGTCAGATGTTGGTTAAAGTCCCGCAACGAC 308821  
 QY 1075 GCAACCTTGTCTAATTAATTCATCAATTTGGTGGCACTTATATGAGTCCCGGTGAC 1134  
 DB 308822 GCAACCTTGTCTAATTAATTCATCAATTTGGTGGCACTTATATGAGTCCCGGTGAC 308881  
 QY 1135 AAACCGAAGAGGTGGGATGACATCAAGTCTCAATGCGCCCTATGGGTGAAGGCTTCA 1194  
 DB 308882 AAACCGAAGAGGTGGGATGACATCAAGTCTCAATGCGCCCTATGGGTGAAGGCTTCA 308941  
 QY 1195 ACCTAATCAATGCGCGTACAGAGGGTTCCAACCGCGAGGGGAGCTAATCTCGAA 1254  
 DB 308942 ACCTAATCAATGCGCGTACAGAGGGTTCCAACCGCGAGGGGAGCTAATCTCGAA 309001  
 QY 1255 AGCGCGTCTGATGTCGCGATCGAGTCTGCAATCTGACTCGGTGAAGTGGAAATCGTACT 1314  
 DB 309002 AGCGCGTCTGATGTCGCGATCGAGTCTGCAATCTGACTCGGTGAAGTGGAAATCGTACT 309061  
 QY 1315 AATCGCGGATCAGCATGTCCCGGTGAATACGTTCCCGGGCTTGTATACACCGCGCTCA 1374  
 DB 309062 AATCGCGGATCAGCATGTCCCGGTGAATACGTTCCCGGGCTTGTATACACCGCGCTCA 309121  
 QY 1375 CACCATGGAGTGGGTTTCAACAGAACAGGTAGTCTAACCCTTAAGAGAGGGCGCTTCCA 1434  
 DB 309122 CACCATGGAGTGGGTTTCAACAGAACAGGTAGTCTAACCCTTAAGAGAGGGCGCTTCCA 309181  
 QY 1435 CGGTGAGATTCAATGACTGGGGTG 1457  
 DB 309182 CGGTGAGATTCAATGACTGGGGTG 309204

RESULT 55  
 ABZ69297  
 ID ABZ69297 standard; DNA; 1400 BP.

XX  
 AC ABZ69297;  
 XX  
 DT 11-AUG-2003 (first entry)  
 XX  
 DE J lividum 16s ribosomal RNA gene fragment #1.  
 XX  
 KM Osteoarthritis; antibacterial; Janthinobacterium; 16s RNA; gene;  
 XX osteopathic; antiarthritic; analgesic; anti-inflammatory; ds.  
 OS Janthinobacterium lividum.  
 PN W02002102384-A1.  
 XX  
 PD 27-DEC-2002.  
 XX  
 PF 17-JUN-2002; 2002WO-GB002771.  
 XX  
 PR 15-JUN-2001; 2001GB-00014672.  
 XX  
 PA (ORTH-) ORTHOGENICS AS.  
 PA (GARD/) GARDNER R.  
 XX  
 PI El-Gewely MR;  
 XX  
 DR WPI; 2003-175199/17.  
 XX  
 PT Use of an antibacterial agent in the manufacture of a medicament for  
 PT treating osteoarthritis.  
 XX  
 PS Example 1; Page 46-48; 89pp; English.  
 XX  
 CC The present invention relates to the use of an antibacterial agent in the  
 CC production of a treatment for osteoarthritis. The bacteria causing  
 CC osteoarthritis is Janthinobacterium lividum. The present sequence is a  
 CC fragment of the 16s RNA coding sequence from J. lividum shown in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 1400 BP; 370 A; 309 C; 429 G; 287 T; 0 U; 5 Other;  
 Query Match 77.1%; Score 1123.2; DB 8; Length 1400;  
 Best Local Similarity 88.8%; Pred. No. 0;  
 Matches 1249; Conservative 0; Mismatches 148; Indels 10; Gaps 3;  
 QY 27 CATGCAAGTCGAACGCGACGACGATGCTTGCATCTGCTGGGAGTGGCGGACGGGTGAG 86  
 DB 1 CATGCAAGTCGAACGCGACGACGGA--GCTTGTCTGTGGGAGTGGCGGAAAGGGGTGAG 58  
 QY 87 TAATGATGGAACCTATCCAGAAAGGGGGTAAACGATCGAAAGATGTGCTAATACCG 146  
 DB 59 TAATATATCGGAACCTATCCAGAAAGGGGGTAAACGATCGAAAGATGTGCTAATACCG 118  
 QY 147 CATATCTCTAAGAGAGAAAGCAGGGGATCGAAAGACCTTGGCTTTGGAGCGGCGAT 206  
 DB 119 CATATCATCTAAGAGAGAAAGTGGGGATCGAAAGACCTATGCTGATGATACCG 178  
 QY 179 ATCTGATTAAGCTATGTTGTTAGGGTAAAGCTTACCAAGGATGATCAGTACGTGTCTG 238  
 DB 207 GTCTGATTAAGCTATGTTGTTAGGGTAAAGCTTACCAAGGATGATCAGTACGTGTCTG 266  
 QY 267 AGAGAACGACCAACCACTTGGGATCGAGACCGGCGCAAGCTCTACCGGAGGACGACAG 326  
 DB 239 AGAGAACGACCAACCACTTGGGATCGAGACCGGCGCAAGCTCTACCGGAGGACGACAG 298  
 QY 327 TGGGAAATTTTGGACATGAGGCGCAAGCTGATCCAGCAATGCGCGGTGAGTGAAGAGG 386  
 DB 299 TGGGAAATTTTGGACATGAGGCGCAAGCTGATCCAGCAATGCGCGGTGAGTGAAGAGG 358  
 QY 387 CCTTGGGTTTGAAGCTCTTTCATGTCAGAGAAAGAAAGTTACGTAATTAATCCGACT 446  
 DB 359 CCTTGGGTTTGAAGCTCTTTCATGTCAGAGAAAGAAAGTTACGTAATTAATCCGACT 418  
 QY 447 CATGACGATTCAGACGAAGAACACCGGCTAATCTAGTCCAGACCGCGGTAAATACG 506



```
Db 419 AATGACGGTACTGTAAGAAATTAAGCACCAGCTAATCACTGCGCAGCAGCAGCCGGGTAAATACG 478
Qy 507 TAGGGTGCAGCGCTTAATCGAATTAATCTGGGCGTAAAGGGTGCAGCGCGCTTTGTAAAG 566
Db 479 TAGGGTGCAGCGCTTAATCGAATTAATCTGGGCGTAAAGGGTGCAGCGCGCTTTGTAAAG 538
Qy 567 TCAGATGTGAATATCCCGGGCTTAACCTGGGAATTCGTTGAATCAAGGCTAGAGT 626
Db 539 TCTGATGTGAATATCCCGGGCTTAACCTGGGAATTCGTTGAATCAAGGCTAGAGT 598
Qy 627 GTGCGAGAGGAGTGAATTCATGTGAGCAGTGAATTCGGTGAAGATATGAGAAAGC 686
Db 599 CTGGCAGAGGGGGGTAGAAATTCACGTGACGAGTGAATTCGTGATATGTGGAGAGAC 658
Qy 687 ATCGATGGGGAAGGAGCGCTCTGGGTAAACACTGACGCTCATGACGAAACGTTGGGGA 746
Db 659 ACCGATGGGGAAGGAGCGCGCTGGGTCAAGATTGACGCTCATGACGAAACGTTGGGGA 718
Qy 747 GCAACAGGATTAAGTACCTCTGTGATGTCACGCGCTTAAACGATGTCAATTAATTTGGG 806
Db 719 GCAACAGGATTAAGTACCTCTGTGATGTCACGCGCTTAAACGATGTCAATTAATTTGGG 778
Qy 807 CCTTATTAAGGCTTGGTAAAGAGCTAACGCGTGAAGTTGACCGCTGGGGAGTACGCTGC 866
Db 779 TCTTATTAAGGCTTGGTAAAGAGCTAACGCGCTGATGACGCGCTGGGGAGTACGCTGC 838
Qy 867 CAAGATTAATACTCAAGAGAAATTGACGCGGAGCCGCAACGCGGTGATTAATGTGATTA 926
Db 839 CAAGATTAATACTCAAGAGAAATTGACGCGGAGCCGCAACGCGGTGATTAATGTGATTA 898
Qy 927 ATTGATGCAAGCGGAAAAAAGCTTACCTTACCTTGAATGAGCGAATTTCTAAGATTA 986
Db 899 ATTGATGCAAGCGGAAAAAAGCTTACCTTACCTTGAATGAGCGAATTTCTAAGATTA 958
Qy 987 GATTAGTGTCT--TCGGGAACGCTAACACAGGTCGTCATGCTGTCAGCTGCTGTC 1043
Db 959 AGGGAGTGTCTGAAAGAGAACCAAGTACACAGGTCGTCATGCTGTCAGCTGCTGTC 1018
Qy 1044 GTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGTCTATTAAATTTGCCATCATTT 1103
Db 1019 GTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGTCTATTAAATTTGCCATCATTT 1073
Qy 1104 GGTGGGCACTTAAATGAGACTGCGGGTACAAACCGAGGAAGGTGGGGAATGACGTCAA 1163
Db 1074 GAAAGGGCACTTAAATGAGACTGCGGGTACAAACCGAGGAAGGTGGGGAATGACGTCAA 1133
Qy 1164 GTTCCTCATGGACCTTAATGGGTAGGGCTTACACGTCATACATATGTAATCAAGAGCGCC 1223
Db 1134 GTTCCTCATGGACCTTAATGGGTAGGGCTTACACGTCATACATATGTAATCAAGAGCGCC 1193
Qy 1224 GCCAACCCCGCAGGGGAGCTAATCTCAGAAAAGCGGCTGTAGTCCGATTCGAGTCTGC 1283
Db 1194 GCCAACCCCGCAGGGGAGCTAATCGAGAAAAGTGTATGTAGTCCGATTTGATGCTGC 1253
Qy 1284 AACTGATCTCCGTGAAGTGGAAATGCTTATGCGGATCAAGCATGTCCGGGTGAATA 1343
Db 1254 AACTGATCTCCGTGAAGTGGAAATGCTTATGCGGATCAAGCATGTCCGGGTGAATA 1313
Qy 1344 CGTTCGCGGGCTTGTACACACGCGCCGTCAACCATGAGAGTGGGATTTTACCAAGATTA 1403
Db 1314 CGTTCGCGGGCTTGTACACACGCGCCGTCAACCATGAGAGTGGGATTTTACCAAGATTA 1373
Qy 1404 GGTAGTCTAACCGTAAGAGAGCGCTT 1430
Db 1374 GGTAGTCTAACCGTAAGAGAGCGCTT 1400
```

```
RESULT 56
ABV72366
ID ABV72366 standard; DNA; 1452 bp.
XX
AC ABV72366;
```

```
XX 16-DEC-2002 (first entry)
DT
XX Nucleotide sequence of a bacterial 16S rRNA gene fragment.
DE
XX 16S rRNA gene; beta proteobacteria; trichloroethylene; soil purification;
KM underground water purification; phenol; ss.
XX Unidentified.
XX JP2002142756-A.
PN
XX 21-MAY-2002.
PD
XX 09-NOV-2000; 2000JP-00341513.
XX 09-NOV-2000; 2000JP-00341513.
PR
XX 09-NOV-2000; 2000JP-00341513.
XX (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.
PA WPI; 2002-676987/73.
DR
XX Novel bacteria belonging to beta proteobacteria for purifying soil or
PT underground water contaminated with phenol and trichloroethylene,
PT decompose trichloroethylene in the presence of phenol under aerobic
PT conditions.
XX Claim 1; Page 5-6; 9pp; Japanese.
PS
XX The present sequence represents a fragment of the 16S rRNA gene of
CC bacteria belonging to beta proteobacteria. The bacteria are capable of
CC decomposing trichloroethylene with high efficiency. The bacteria are
CC useful for purifying soil or underground water contaminated with phenol
CC and trichloroethylene
XX
SQ Sequence 1452 bp; 375 A; 338 C; 448 G; 291 T; 0 U; 0 Other;
Query Match 77.1%; Score 1123; DB 6; Length 1452;
Best Local Similarity 87.5%; Pred. No. 0; Mismatches 175; Indels 6; Gaps 3;
Matches 1264; Conservative 0;
Qy 16 CATGCTTACACATGCAAGTGAACGCGCAGCAGGATGCTTGCATCTGGTGGCGAGTGGC 75
Db 1 CATGCTTACACATGCAAGTGAACGCGCAGCAGGATGCTTGCATCTGGTGGCGAGTGGC 58
Qy 76 GACGCGGTGAGTATGATCGGAACGTATCCAGAAAGAGGGGGTAAACGATCGCAAGATG 135
Db 59 GACGCGGTGAGTATGATCGGAACGTATCCAGAAAGAGGGGGTAAACGATCGCAAGATG 118
Qy 136 TGCTAATACCGCATATATCTCAAGAGGAGGAGGAGATGGAAGACCTTGCGCTTTTG 195
Db 119 TGCTAATACCGCATATATCTCAAGAGGAGGAGGAGATGGAAGACCTTGCGCGAATG 178
Qy 196 GAGCGCGCGATGTGATTAAGTGTGGTGGGTTAAAGGCTTACCAAGGCGAGATCAG 255
Db 179 GAGCGCGCGATGTGATTAAGTGTGGTGGGTTAAAGGCTTACCAAGGCGATGATG 238
Qy 256 TAGTTGTCTGAGAGGAGCAGCAGCAGCAGCTGGGACTGAGACAGCGGCCAGATCTTA 315
Db 239 TAGTTGTCTGAGAGGAGCAGCAGCAGCAGCTGGGACTGAGACAGCGGCCAGATCTTA 298
Qy 316 GAGGCGAGCATGGGGAAATTTTGAACAATGGGCCCAAGCCGTATCCAGCAATGCGCGTG 375
Db 299 GAGGCGAGCATGGGGAAATTTTGAACAATGGGCCCAAGCCGTATCCAGCAATGCGCGTG 358
Qy 376 AGTGAAGAGGCTTGGGTTGTAAAGCTTTTCAAGTGCAGAAAGAAAGTTAGCGTAA 435
Db 359 CAGATGAAGGCTTGGGTTGTAAAGCTTTTCAAGTGCAGAAAGAAAGTTAGCGTAA 418
Qy 436 TAAATGTAATCATGACGATTCAGACAGAAAGACCGGCTAACTACGTCCAGACACC 495
Db 419 TACAGAAAGGCTTAATGACGATCCGTAAGAAATTAAGACCGGCTAACTACGTCCAGACACC 478
```



QY	496	GC	CGGTAA	TAC	TAC	TAG	TGG	TGC	AAC	CGCTTA	TAT	GG	GA	TTC	CTG	GGG	GGT	TAA	AGG	GTG	CG	A	GC	GC	555
Db	479	G	CGGTAA	T	T	A	C	T	A	C	T	A	A	G	T	T	A	C	T	G	G	G	T	A	538
QY	556	GG	CTTT	G	AG	T	CA	G	AT	GT	CA	AA	T	CC	CC	GGG	CT	T	AA	C	CT	GG	GA	T	615
Db	539	G	GT	T	A	T	A	T	A	G	A	C	A	G	T	T	G	A	A	T	C	C	C	C	598
QY	616	A	AG	CT	A	G	A	G	T	GT	GC	A	A	G	G	A	G	GT	GA	A	T	T	C	A	675
Db	599	A	T	A	C	T	A	G	A	G	T	A	G	G	G	G	A	T	T	C	C	G	G	T	658
QY	676	T	A	T	A	G	A	A	C	A	T	G	A	T	G	C	A	A	G	G	C	A	C	T	735
Db	659	T	G	C	G	A	G	G	A	C	A	C	G	A	T	G	C	A	A	G	G	C	A	T	718
QY	736	A	A	G	G	T	G	G	G	A	G	A	A	C	A	A	G	A	T	T	A	G	T	A	795
Db	719	A	A	G	G	T	G	G	G	A	C	A	A	G	A	T	T	A	G	T	A	G	T	A	778
QY	796	T	A	G	T	G	T	G	G	C	T	T	A	T	T	A	G	C	T	T	G	T	A	C	855
Db	779	T	G	T	T	G	T	T	G	G	T	C	T	T	A	C	T	G	A	C	T	G	A	C	838
QY	856	G	A	G	T	A	C	G	T	C	C	G	A	A	T	T	A	A	A	C	T	C	A	A	915
Db	839	G	A	G	T	A	C	G	T	C	C	G	A	A	T	T	A	A	A	C	T	C	A	A	898
QY	916	T	A	T	A	T	G	A	T	T	A	T	T	G	A	T	T	G	A	A	A	A	A	A	975
Db	899	G	A	T	A	T	T	A	T	T	A	T	T	G	A	T	T	G	A	A	A	A	A	A	958
QY	976	T	T	C	T	A	G	A	T	A	G	T	A	T	A	G	T	G	C	T	---	T	C	G	1032
Db	959	C	C	C	A	G	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	1018
QY	1033	C	A	G	C	T	C	T	G	T	C	G	A	T	G	T	T	A	A	G	T	T	A	A	1092
Db	1019	C	A	G	C	T	C	T	G	T	C	G	A	T	G	T	T	A	A	G	T	T	A	A	1078
QY	1093	T	G	C	A	T	C	A	T	T	G	T	T	G	G	C	A	C	T	T	T	A	T	A	1152
Db	1079	T	G	C	T	A	C	A	T	T	G	T	T	G	G	C	A	C	T	T	T	A	T	A	1137
QY	1153	G	A	T	A	C	G	T	C	A	A	G	T	C	C	T	T	A	T	G	G	G	C	T	1212
Db	1138	G	A	T	A	C	G	T	C	A	A	G	T	C	C	T	T	A	T	G	G	G	C	T	1197
QY	1213	T	A	C	A	G	A	G	G	T	T	G	C	A	C	C	G	A	G	G	G	A	G	A	1272
Db	1198	T	A	C	A	G	A	G	G	T	T	G	C	A	C	C	G	A	G	G	G	A	G	A	1257
QY	1273	T	G	G	A	G	T	C	T	G	C	A	A	G	T	C	C	G	T	A	A	G	T	C	1332
Db	1258	T	G	C	A	G	T	C	T	G	C	A	A	G	T	C	C	G	T	A	A	G	T	C	1

AC	ADB61691;
DT	04-DEC-2003 (first entry)
DE	16S rRNA of Burkholderia pseudomallei DNA sequence.
KW	enriching mRNA; high quality bacterial mRNA; bacterial gene expression;
KW	poly-A tail; mRNA purification; oligo-dT capture;
KM	prokaryote mRNA purification; bridging oligonucleotide; targeting region;
KM	capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;
KW	eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;
KX	28S eukaryotic RNA; bridging oligonucleotide; ds.
OS	Burkholderia pseudomallei.
PN	WO2003054162-A2.
PD	03-JUL-2003.
XX	19-DEC-2002; 2002MO-US041014.
PF	20-DEC-2001; 2001US-00029397.
PR	(AMBI-) AMBION INC.
PA	Murphy GL, Whitley JP;
PI	WPI: 2003-663255/62.
DR	Depleting or isolating targeted nucleic acids e.g. rRNA, involves using a
PT	bridging oligonucleotide comprising bridging region and a targeting
PT	region complementary to a targeted nucleic acid, and a capture
XX	oligonucleotide.
XX	Claim 4; Page 175; 208pp; English.
XX	This invention relates to a novel method for isolating, depleting or
CC	separating a targeted nucleic acid, such as rRNA, from a sample
CC	comprising targeted and non-targeted nucleic acids. It effects a way of
CC	enriching for non-targeted nucleic acids such as rRNAs. Isolating
CC	sufficient quantities of high quality bacterial mRNA is a demanding
CC	process which impedes analysis of bacterial gene expression in the
CC	presence of host cells. A small percentage of bacterial mRNAs may be poly
CC	-A tailed, but these are targeted for degradation and tend to be
CC	unstable. As a result, the commonly employed method for mRNA purification
CC	with eukaryotic cells, oligo-dT capture, is ineffective. The present
CC	invention provides an alternative, more suitable method for mRNA
CC	purification from prokaryotes. The method of the invention comprises the
CC	incubation of a sample with a bridging oligonucleotide (containing a
CC	targeting region) and subsequently incubating with a capture
CC	oligonucleotide allowing the isolation of the target from the sample. The
CC	method is useful for depleting or isolating targeted nucleic acid, for
CC	example rRNA such as prokaryotic 16S or prokaryotic 23S, eukaryotic 17S
CC	or 18S or eukaryotic 28S rRNA, from a sample. The rRNA sequence may
CC	comprise any one of 64 fully defined sequences as given in the
CC	specification. The present sequence is that of a DNA sequence which
CC	represents the sequence of 16S rRNA of Burkholderia pseudomallei related
CC	to the invention.
XX	Sequence 1610 BP; 394 A; 383 C; 514 G; 319 T; 0 U; 0 Other;
SQ	
Query Match	77.0%; Score 1122; DB 10; Length 1610;
Best Local Similarity	87.0%; Pred. No. 0;
Matches 1270; Conservative	0; Mismatches 180; Indels 10; Gaps 3
QY	1 ATTGAACGCTGGCGGCATGCTTTAACATGCAATGCCAAAGTCGACGACAGGAGCTTGCAAT 60 
DB	82 ATTTGAACGCTGGCGGCATGCTTTAACATGCAATGCCAAAGTCGACGACAGG--GCTTCGACC 139 
QY	61 CTGTGTGGCAGTAGTGGGAGACGGGTGATGATGATCGAAGCATATCAGAAGAGGAGGATTA 120 
DB	140 CTGTGTGGCAGTAGTGGGAGACGGGTGATGATGATCGAAGCATATCCTTAAGTGGGGATTA 199 

QY	121	ACGCATCGAAGAATGTGCTAATACCGCATATACTCTTAAGAGAGAAAGCAAGGGATATCGAAA	180
Db	200	GCCCCCGCCAAAGCCCAATTAAATACCGCATACGATTCGAGGAATGAAGGCGGGGACCTTTCG	259
QY	181	GACCTTGGGCTTTTGGAGCGGCCCAATGTCTGATTTAGCTAGTTGGTGGGGTAAAGGCTTAC	240
Db	260	GGCCTTCGGGCTATAGGGTTGGCCGATGGCTGATTTAGCTAATGTTGGTGGGGTAAAGGCTTAC	319
QY	241	CAAGGCGACGATCACTAGTTGGTCTGAGAGACGACCAAGCCACTGGGGACTGAGACAG	300
Db	320	CAAGGCGACGATCACTAGTTGGTCTGAGAGACGACCAAGCCACTGGGGACTGAGACAG	379
QY	301	GCCCGACATCTCTTACGGGAGGCGACAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC	360
Db	380	GCCCGACATCTCTTACGGGAGGCGACAGTGGGAAATTTTGGACAAATGGGCGCAAGCTGATC	439
QY	361	CAGCAATGCGCGGTAGTGAAGAAGAGCCCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAGA	420
Db	440	CAGCAATGCGCGGTAGTGAAGAAGAGCCCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAGA	499
QY	421	AAAGGTTACGGTAAATTAATCTGACTACACGATCGACAGAGAAAGCACCGGCTTAC	480
Db	500	AATCATTTCTGGCTAATATACCGGAGTGGATGACGATACCGGAAGAAATTAAGCACCGGCTTAC	559
QY	481	TACGTGCGACAGCGCCCGGTAAATACGTAGGGTGGCAAGCTTAATGGAATTACTGGGGCT	540
Db	560	TACGTGCGACAGCGCCCGGTAAATACGTAGGGTGGCAAGCTTAATGGAATTACTGGGGCT	619
QY	541	AAAGGATGGCGAGCGGGCTTTGTAAATGTCAGATGTGAATCCCGGGCTTAACTGGGAAT	600
Db	620	AAAGGATGGCGAGCGGGCTTTGTAAATGTCAGATGTGAATCCCGGGCTTAACTGGGAAT	679
QY	601	TGCGTTTGAATACTACAAGGCTTAGATGTGGCAGAGGAGGTGAATTCATGTGACAG	660
Db	680	TGCATTTGGATGCTGGCAGGCTTAGATGTGGCAGAGGAGGTGAATTCATGTGACAG	739
QY	661	TGAATGCGTAGAGATATGGAAGAACATGATGGGAGGACGCTCTCGGGTTAACT	720
Db	740	TGAATGCGTAGAGATATGGAAGAACATGATGGGAGGACGCTCTCGGGCTTAATCT	799
QY	721	GACGTCATGCACGGAAGCGTGGGAGGAAACAGATTAAGATACCCTGTAGTCCAGGCC	780
Db	800	GACGTCATGCACGGAAGCGTGGGAGGAAACAGATTAAGATACCCTGTAGTCCAGGCC	859
QY	781	CTAAACGATGCACTAGTTGTTGGGCTTATTAAGCTTGGTAAAGAACTTACCGCTGA	840
Db	860	CTAAACGATGCACTAGTTGTTGGGATTAATTTCTTAACTGTAAGCTTACCGCGGA	919
QY	841	AGTTGACCGCTGGGAGAGTACGGTTCGACAGATTTAAACTCAAGGAATTGACGGGAGCC	900
Db	920	AGTTGACCGCTGGGAGAGTACGGTTCGACAGATTTAAACTCAAGGAATTGACGGGAGCC	979
QY	901	GCAAGAGGGTGGGATTAATGTGGAATTAATTCGATGGAACGCGAATAACTTACCTACCCCTT	960
Db	980	GCAAGAGGGTGGGATTAATGTGGAATTAATTCGATGGAACGCGAATAACTTACCTACCCCTT	1039
QY	961	GACATGTAGCGAATTTTCTAGAGATAGATTAGTGTCTCTCGGAGACCTTAAACAGGTG	1014
Db	1040	GACATGTAGCGAAGCCCGATGAGATTTGGGCTCTCCAGAAAGAAACCGGCGCACAGTGTG	1099
QY	1018	CTGCATGGCTGTGCTCAGCTCGTGTCTGTAGATGTGTGGGTTAAAGTCCGCGCAAGCGCA	1077
Db	1100	CTGCATGGCTGTGCTCAGCTCGTGTCTGTAGATGTGTGGGTTAAAGTCCGCGCAAGCGCGCA	1156
QY	1078	ACCTTTGTCATTAATTTGTCATATTTGGTGGGCACTTAAATGAAACTGCGGTGACAA	1137
Db	1160	ACCTTTGTCATTAATTTGTCATATTTGGTGGGCACTTAAAGGAATGCGGTGACAA	1214
QY	1138	CCGAGAGAAAGTGGGAGATGACGTCAAGTTCCTATGAGCCCTTAATGGGTAGGGCTTACACG	1197
Db	1215	CCGAGAGAAAGTGGGAGATGACGTCAAGTTCCTATGAGCCCTTAATGGGTAGGGCTTACACG	1274
QY	1198	TAAATCAATGGCGCTTACAGAGGGTGGCCAAACCGCGAGGGGAGCTAATCTCAGAAAGC	1257

Accession	Sequence	Length
Db	1275 TCATTCMAATGGTCCGGAAACAGAGGGGTCCGCAACCCGGCAGGGGGAGGCCAATCCACGAAAAAC	1334
QY	1258 GCGTCGTATGTCGGATCGGAAGTCTTGCAACTTCGACTCCGTGAAGTCGGAATGGCTATGTAAT	1317
Db	1335 CGATCGTAGTCCGGAATTCGACTCTTGCAACTTCGAGTGAAGTGAAGCTGGAAATGCTATGTAAT	1399
QY	1318 CGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTTGTACACACCGCCCGTCACAC	1377
Db	1395 CGCGGATCAGCATGTCGCGGTGAATACCTTCCCGGGTCTTTGTACACACCGCCCGTCACAC	1454
QY	1378 CATGGGAATGGGTTTACACGAAGAGGTAGTCTTAACCGTTAGGAGGGCGCTTGCCACGG	1437
Db	1455 CATGGGAATGGGTTTACACGAAGTGGCTTACTTAAACCGCAAGAGAGCGGTACACACGG	1514
QY	1438 TGAATTTCATGACTGGGGTG	1457
Db	1515 TAGGATTCATGACTGGGGTG	1534

## RESULT 58

ID AAS11025 standard; DNA; 1544 BP.

AC AAS11025;

DT 06-AUG-2003 (revised)

XX

XX

KW food grain supplement; livestock; poultry; therapeutic; drugs

OS *Neisseria gonorrhoeae*.

PN W0200142457-A2.

PD 14-JUN-2001

PF 29-NOV-2000; 2000WO-US042391.

PR 29-NOV-1999; 99US-0168150P.

PA (AVIB-) AVI BIOPHARMA INC.

PI Iversen PL;

DR WPI; 2001-457295/49.

PT Antibacterial compound, useful for treating bacterial infections and as

PT oligonucleotides complementary to bacterial 16S and 23S rRNA

PS Disclosure; Page; 62pp; English.

CC AAS11021,AAS11034 represent the coding sequences of bacterial 16S-  
CC ribosomal RNA (rRNA) genes. The sequences were used to design anti-  
CC bacterial compounds comprising substantially unchanged antisense  
CC oligomers containing 8-40 nucleotide subunits, including a targeting  
CC nucleic acid sequence at least 10 nucleotides in length which is  
CC complementary to a bacterial 16S or 23S rRNA nucleic acid sequence. The  
CC antisense oligomers are used for treating a bacterial infection in a  
CC human or a mammalian animal produced by *Escherichia coli*, *Salmonella*  
CC typhimurium, *Pseudomonas aeruginosa*, *Vibrio cholerae*, *Neisseria*  
CC gonorrhoea, *Helicobacter pylori*, *Bartonella henselae*, *Haemophilus*  
CC influenza, *Shigella dysenteriae*, *Staphylococcus aureus*, *Mycobacterium*  
CC tuberculosis, *Streptococcus pneumoniae*, *Treponema pallidum* and *Chlamydia*  
CC trachomatis. The antibacterial compound may be used as a food grain  
CC supplement in livestock and poultry food composition. Note: The present  
CC sequence is not shown in the specification but has been accessed from  
CC Genbank using the appropriate accession number given in the  
CC specification. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 1544 BP; 380 A; 357 C; 492 G; 315 T; 0 U; 0 Other;  
 SQ Query Match 76.9%; Score 1119.8; DB 5; Length 1544;  
 Best Local Similarity 87.1%; Pred. No. 0; Mismatches 182; Indels 6; Gaps 4;  
 Matches 1275; Conservative 0;

QY 1 ATTGAACGCTGGCGGATCTTTACATGCAAGTCGAACGGCAGAC--GGATGCTTGC 58  
 DB 29 ATTGAACGCTGGCGGATCTTTACATGCAAGTCGAACGGCAGACGGAGCTTGC 88  
 QY 59 ATCT--GGTGGAGTGGCGGACGGGTGATTAATGCATCGAAACGTATCAAGAAAGGGG 116  
 DB 89 TTCTGGGTGGCGAGTGGCGAAACGGGTGATTAACATATCGAAACGTATCGGGTAAAGGGG 148  
 QY 117 GGTAAACGATCGAAAGATGTCTAATAGCGCATATATCTTAAGGAGAAAGCAGGGGATC 176  
 DB 149 GATTAACGTATCGAAAGATGATTAATAGCGCATATCTTGAAGAGAAAGCAGGGGATC 208  
 QY 177 GAAAGACCTTGGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTTGTGGGTAAAGGC 236  
 DB 209 TTCCGGGCTTTGGCGCTATCCGAGGCGCATATCTGATTAGCTGTTGGCGGGTAAAGGC 268  
 QY 237 CTACCAAGGCGACGATCAAGTGTGTCTGAGAGAGCAACGCCACACTGGGACTGAGA 296  
 DB 269 CCACCAAGGCGACGATCAAGTGTGTCTGAGAGAGATCCGCCACACTGGGACTGAGA 328  
 QY 297 CACGGCCGACATCTCTTAACGGAGAGCAGTGGGGAATTTTGGACATGGGCGCAAGCCT 356  
 DB 329 CACGGCCGACATCTCTTAACGGAGAGCAGTGGGGAATTTTGGACATGGGCGCAAGCCT 388  
 QY 357 GATCCAGCATGCGCGGTGATGAAGAAAGCCTTCGGGTGTAAAGCTCTCATGTCGAG 416  
 DB 389 GATCCAGCATGCGCGGTGATGAAGAAAGCCTTCGGGTGTAAAGCTCTCATGTCGAG 448  
 QY 417 AAGAAAGGTTACGTTAATATCTGATCATGACGCTATGACAGAGAAAGCACCAGC 476  
 DB 449 AAGAAAGGTTACGTTAATATCTGACGCTATGACAGAGAAAGCACCAGC 508  
 QY 477 TAACTACGTGCCAGACCGCGGTAAATAGTAGGGTGAACGGTTATCGGAATTACTGG 536  
 DB 509 TAACTACGTGCCAGACCGCGGTAAATAGTAGGGTGAACGGTTATCGGAATTACTGG 568  
 QY 537 GCGTAAAGGGTGGCAGACGGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACCTGG 596  
 DB 569 GCGTAAAGGGTGGCAGACGGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACCTGG 628  
 QY 597 GAATTCGTTTGAACCTACAAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGA 656  
 DB 629 GAATTCGTTTGAACCTACAAAGGCTAGAGTGTGTGAGAGGAGGTGGAATTCACGTGA 688  
 QY 657 GCAGTGAATGCGTAGAGATATGGAAGAACATGATGCGCAAGGCAAGCCTTCGGGTGA 716  
 DB 689 GCAGTGAATGCGTAGAGATATGGAAGAACATGATGCGCAAGGCAAGCCTTCGGGTGA 748  
 QY 717 CACTGACGCTCATGACGAAAGCGTGGGAGCAAAACGATTAAGTACCTCGTAGTCA 776  
 DB 749 CACTGACGCTCATGACGAAAGCGTGGGAGCAAAACGATTAAGTACCTCGTAGTCA 808  
 QY 777 CGCCCTAAACGATGTCAACTAGTGTGGCTTATTA--GGCTTGTAAACGAACTAACG 835  
 DB 809 CGCCCTAAACGATGTCAACTAGTGTGGCTTATTA--GGCTTGTAAACGAACTAACG 868  
 QY 836 CGTGAAGTTGACCGCTGGGAGATACGCTGGCAAGATTAATACTCAAGGAATTGACGG 895  
 DB 869 CGTGAAGTTGACCGCTGGGAGATACGCTGGCAAGATTAATACTCAAGGAATTGACGG 928  
 QY 896 GACCCGCAAGCGGTGATATATGATTAATTCATGCAACGCGAAGAACTTACCTTA 955  
 DB 929 GACCCGCAAGCGGTGATATATGATTAATTCATGCAACGCGAAGAACTTACCTTA 988  
 QY 956 CCCTTGAATGTAGCGGAATTTTCTAGAGATAGATTAGTG--CTTCGGGAACGCTTAACAG 1014

DB 989 GTTTTGAATGTGGGAATCTCCGAGACGAGAGAGTGCCTTCGGAGCCGTAAACAG 1048  
 QY 1015 GTGCTGATGCGCTGTGCTGACCTGTGTCTGGAATGTTGGGTTAATGTCGCCAACGAGC 1074  
 DB 1049 GTGCTGATGCGCTGTGCTGACCTGTGTCTGGAATGTTGGGTTAATGTCGCCAACGAGC 1108  
 QY 1075 GCAACCTTGTCAATTAATTTGATCATATTTGGTGGGCACTTTAATGAGACTGCCGTGAC 1134  
 DB 1109 GCAACCTTGTCAATTAATTTGATCATATTTGGTGGGCACTTTAATGAGACTGCCGTGAC 1168  
 QY 1135 AAACCGAGAAAGGTGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGCTTAC 1194  
 DB 1169 AAGCCGAGAAAGGTGGGATGACGTCAAGTCTCATGAGCCCTTAATGAGAGGCTTAC 1228  
 QY 1195 ACGTAAATCAATGCGCGCTACAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAA 1254  
 DB 1229 ACGTAAATCAATGCGCGCTACAGAGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAA 1288  
 QY 1255 AGCGGCTGTAGTCCGGATCCGAGTCTGCACTCGACTCCGTAAGTCGGAATCGTAGT 1314  
 DB 1289 AACCAGTGTAGTCCGATGATGCACTTGCATGAGTGCATGAGTCCGAATCGCTAGT 1348  
 QY 1315 AATCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCA 1374  
 DB 1349 AATCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCA 1408  
 QY 1375 CACATGGGAGTGGGTTTCAACGAGAGAGTGTCTAACCTTAAGAGGGGCTTTGCCA 1434  
 DB 1409 CACATGGGAGTGGGAGTATCAAGAGTGTGTAAACCGCAAGAGTCCGCTTAACA 1468  
 QY 1435 CGGTAGATTTCACTGCTGGGCTG 1457  
 DB 1469 CGGTAGTCTTCACTGCTGGGCTG 1491

RESULT 59  
 ADB61692  
 ID ADB61692 standard; DNA; 1544 BP.  
 AC XX  
 AC ADB61692;  
 DT XX  
 DT 04-DEC-2003 (first entry)  
 XX XX  
 XX 16S rRNA of *Neisseria gonorrhoeae* DNA sequence.  
 DE XX  
 DE enriching mRNA; high quality bacterial mRNA; bacterial gene expression;  
 KW poly-A tail; mRNA purification; oligo-dT capture;  
 KW prokaryote mRNA purification; bridging oligonucleotide; targeting region;  
 KW capture oligonucleotide; prokaryotic 16S rRNA; prokaryotic 23S rRNA;  
 KW eukaryotic 17S rRNA; eukaryotic 18S rRNA; eukaryotic 28S rRNA;  
 KW 28S eukaryotic rRNA bridging oligonucleotide; ds.  
 XX XX  
 OS *Neisseria gonorrhoeae*.  
 PN WO2003054162-A2.  
 PD 03-JUL-2003.  
 PD 19-DEC-2002; 2002WO-US041014.  
 PF 20-DEC-2001; 2001US-00029397.  
 XX XX  
 XX (AMBI-) AMBION INC.  
 PA Murphy GL, Whitley JP;  
 XX PI  
 XX MPI; 2003-663255/62.  
 DR  
 XX  
 PT Depleting or isolating targeted nucleic acid e.g. rRNA, involves using a  
 PT bridging oligonucleotide comprising bridging region and a targeting  
 PT region complementary to a targeted nucleic acid, and a capture  
 PT oligonucleotide.  
 XX XX



XX JP11056370-A.  
XX 02-MAR-1999.  
XX 19-AUG-1997; 97JP-00236452.  
XX 19-AUG-1997; 97JP-00236452.  
XX (TOYOTA JIDOSHA KK.  
XX (TOYOTA CHUO KENKYUSHO KK.  
XX WPI; 1999-222383/19.  
XX Detection of a trichloroethylene-decomposing microbe - using a medium  
XX containing the chemical.  
XX Claim 1; Page 11; 17pp; Japanese.  
XX  
XX The present sequence represents a 16S rDNA of a trichloroethylene-  
XX decomposing microbe, Burkholderia sp. N16-1 (FERM BP-5504). The invention  
XX relates to a method for the detection of a trichloroethylene-decomposing  
XX microbe, Burkholderia sp. N16-1 (FERM BP-5504) in which it is selectively  
XX detected or counted by a medium containing the chemical by using the  
XX chemical-resistant strain; It also provides a mutant of the  
XX trichloroethylene-decomposing microbe, Burkholderia sp. N16-1 (FERM BP-  
XX 5504) having chemical resistance and maintaining trichloroethylene-  
XX decomposing activity  
XX  
XX Sequence 1477 BP; 372 A; 346 C; 474 G; 285 T; 0 U; 0 Other:  
SQ  
Query Match 76.8%; Score 1119.4; DB 2; Length 1477;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 12/4; Conservative 0; Mismatches 171; Indels 12; Gaps 4;  
QY 6 ACGTGGCGCATGCTTTACATGCAAGTCGAGCGGACGACGATGCTTGATCGTGGT 65  
DB 4 ACGTGGCGCATGCTTTACATGCAAGTCGAGCGGACGACGATGCTTGATCGTGGT 61  
QY 66 GCGGATGGCGGACGGGTGATGATGCAATCGGAAGTATCCAGAGAGGGGGTTAAGCA 125  
DB 62 GCGGATGGCGGACGGGTGATGATGCAATCGGAAGTATCCAGAGAGGGGGTTAAGCA 121  
QY 126 TCGAAGATGCTGCTAATACCGCATATCTTAAGAGAGAAAGCAGGCGATC--GAAAGAC 183  
DB 122 GCGAAGATGCTGCTAATACCGCATATCTTAAGAGAGAAAGCAGGCGATC--GAAAGAC 181  
QY 184 CTGGCGCTTTGAGAGCGCGCATGCTGATTAAGTCTGATGCTGATGCTGATGCTGATGCT 243  
DB 182 CTGGCGCTTTGAGAGCGCGCATGCTGATTAAGTCTGATGCTGATGCTGATGCTGATGCT 241  
QY 244 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303  
DB 242 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301  
QY 304 CAGACTCTTACGGGAGGACAGATGAGGAAATTTTGAACAATGAGGCGACCTGATCCAG 363  
DB 302 CAGACTCTTACGGGAGGACAGATGAGGAAATTTTGAACAATGAGGCGACCTGATCCAG 361  
QY 364 CAATGCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTCTTTGATGAGAGAGAAA 423  
DB 362 CAATGCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTCTTTGATGAGAGAGAAA 421  
QY 424 GGTACGCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 483  
DB 422 CTTGCTGCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 481  
QY 484 GTGCGACAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 543  
DB 482 GTGCGACAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541  
QY 544 GGGTGGCAGGCGGCTTTGTAAGTCAAGATGTAATCCCGGCTTAAGTGGAGATGCG 603

DB 542 GCGTGGCAGGCGGCTTTGTAAGTGAATCCCGGCTTAAGTGGAGAACTGC 601  
QY 604 GTTGAATCTCAAGGCTAAGTGTGTCAGAGAGGAGGAAATTCATGTTAGACATGA 663  
DB 602 ATTTGTGACTGCGGGCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 661  
QY 664 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 723  
DB 662 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721  
QY 724 GCTCATGACAGAAACGCTGGGAGCAAAAGATTTAGATACCTGTTAGTCCAGCCCTTA 783  
DB 722 GCTCATGACAGAAACGCTGGGAGCAAAAGATTTAGATACCTGTTAGTCCAGCCCTTA 781  
QY 784 AACGATGCACTAATGTTGTTGCTTTATGAGCTTGTGTAACGAAGCTTAAGCGGTAAGT 843  
DB 782 AACGATGCACTAATGTTGTTGCTTTATGAGCTTGTGTAACGAAGCTTAAGCGGTAAGT 841  
QY 844 TGACCGGCTGGGAGTACGATGCGCAAGATTAATACTCAAGAAATGACGGGACCCGCA 903  
DB 842 TGACCGGCTGGGAGTACGATGCGCAAGATTAATACTCAAGAAATGACGGGACCCGCA 901  
QY 904 CAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAAACTTAACCTTACCTTAC 963  
DB 902 CAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAAACTTAACCTTACCTTAC 961  
QY 964 ATGTAGCAATTTCTAGATAGATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 962 ATGTAGCAATTTCTAGATAGATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1021  
QY 1021 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
DB 1022 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1081  
QY 1082 CTG 1136  
DB 1141 GAGGAAGTGGGAGTACGCTCAAGTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
QY 1137 GAGGAAGTGGGAGTACGCTCAAGTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1196  
DB 1201 TACATGCGGCTACAGAGGCTTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
QY 1197 TACATGCGGCTACAGAGGCTTGGCAACCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1256  
DB 1261 TCGTATGCTGCGATGCGAGTCTGCAACTGCACTGCTGCAAGTCTGCAAGTCTGCAAGT 1320  
QY 1257 TCGTATGCTGCGATGCGAGTCTGCAACTGCACTGCTGCAAGTCTGCAAGTCTGCAAGT 1316  
DB 1321 GGATCAGCATGTCGGGTGAATACGTTCCGGGCTTTGTAACAACCGCCGTCACACCAT 1380  
QY 1317 GGATCAGCATGTCGGGTGAATACGTTCCGGGCTTTGTAACAACCGCCGTCACACCAT 1376  
DB 1381 GGGAGTGGGTTTCAACGAAGAGATGATTAACCTTAAGAGAGGGGCTTGGCAAGGTA 1440  
QY 1377 GGGAGTGGGTTTCAACGAAGAGATGATTAACCTTAAGAGAGGGGCTTGGCAAGGTA 1436  
DB 1441 GATTCATGAAGGAGG 1457  
QY 1437 GATTCATGAAGGAGG 1453

RESULT 61  
AAT18765  
ID AAT18765 standard; rRNA; 1536 BP.  
XX  
XX AAT18765;  
AC  
XX 16-OCT-2003 (revised)  
DT 05-JUL-1996 (first entry)  
XX  
XX Pseudomonas testosteroni 16S ribosomal RNA.

XX XX Atrazine; pesticide degradation; soil decontamination; bioremediation;  
KW s-triazine; herbicide; ss.  
XX Comamonas testosteroni.  
OS  
FH Key Location/Qualifiers  
FT misc\_difference 46  
FT /tag= a  
FT /note= "base 46 is identified as 'n'"  
XX  
XX US508193-A.  
XX  
XX 16-APR-1996.  
XX  
XX 31-AUG-1993; 93US-00114695.  
XX  
XX 31-AUG-1993; 93US-00114695.  
XX  
XX (MINU ) UNITV MINNESOTA.  
XX  
XX Wackett LP, Mandelbaum RT;  
XX  
XX MPI; 1996-208726/21.  
XX  
XX Biologically pure culture of atrazine-degrading *Pseudomonas* - useful to  
XX detoxify atrazine, e.g. in soil, at a wide variety of concns.  
XX  
XX Example 2; Col 33-36; 34p; English.  
XX  
XX Novel bacterial strain ADP (ATCC 55464), isolated from atrazine-  
XX contaminated soil, is capable of degrading s-triazine cpds, including  
XX atrazine. In an attempt to identify the strain, the 16S ribosomal RNA  
XX sequence (AAT18760) was compared to that of *Escherichia coli* (AAT18759),  
XX *Pseudomonas citromellolis* ATCC 13674 (AAT18761-63), *Pseudomonas*  
XX *aeruginosa* (AAT18764), *Pseudomonas testosteroni* (AAT18765) and  
XX *Pseudomonas cepacia* (AAT18766). It was concluded that ADP is closely  
XX related to, but distinct from, *P. citromellolis* and *P. aeruginosa*.  
XX (Updated on 16-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 1536 BP; 384 A; 355 C; 482 G; 0 T; 314 U; 1 Other;  
Query Match 76.7%; Score 1118.2; DB 2; Length 1536;  
Best Local Similarity 70.0%; Pred. No. 0;  
Matches 1022; Conservative 252; Mismatches 178; Indels 9; Gaps 4;  
QY 1 ATTGAAGCGTGGCGGAT-GCTTTACATGCAGATCGAAGCGAGCAGATGCTTGA 59  
DB 28 AUGCAGCGUGGCGGCAUNCGCUUACAGCAUGCAUGCAACGUAACAGUCUCG--- 84  
QY 60 TCTGTGGCGAGTGGCGGACGGGTGATGATGCATCGAAGCTATCCAGAGAAGGGGGT 119  
DB 85 -AUGCAGCGUGGCGGACGGGUGAGUAUACAUAGGAACGUGCUAUAUGGGGAGU 143  
QY 120 AACGCATCGAAGATGCTATATCCGCATATCTTTAGAGAGAAAGAGGGAGTCGA 179  
DB 144 AACUACUCGAAAGAGUAGCUAUAACCGCAUGAGAUUACGUAAGAAAGAGGGACUUC 203  
QY 180 AGACCTTGGCTTTTGGAGCGGCGCATGCTGATTAAGTGAAGTGGGGTAAAGCCTA 239  
DB 204 GGGCCUUGUACUAGACCGGCUAUGGCAUAUUAUGUAGUUGGGGUAAGGCUUA 263  
QY 240 CCAAGCGCAGATCACTAGTGGTCTGAGAGAGCAGCAAGCCCACTGGGACTGAGAC 299  
DB 264 CCAAGCGCUGAUCUAGUCUGUUGUAGAGAGCAGCAAGCCCACTGGGACTGAGAC 323  
QY 300 GGGCCAGATCTCTTACGGAGAGCAGAGTGGGAAATTTTGAAGATGGGGGCAAGCCTG 359  
DB 324 GGGCCAGACUCUACGGAGAGCAGAGUGGGAAUUAUUGGAAAGGAAAGGCUUAG 383  
QY 360 CCAGCAATCGCGGTGAGAGAAAGCGCTTCGGGTTGTAAGCTTTCACTCGAAG 419  
DB 384 CAGCAUAGCCGCGUGCAGAGUAAGGCCUCCGGGUGUAACUGUUUUGUACGGAA 443

QY 420 AAAAGTTACGGTAAATTAATGCTGACTCATGACGGTATCCAGAGAAGACCGGCTAA 479  
DB 444 AAAAGCCUGGGCUAUAUUVCCCGGUGUACGAGUACGUAAGAUUAAGCAGCGCUA 503  
QY 480 CTACGTGCACAGACCGCGGTAATATCGTAGGGTCAAGCGCTTAATCGGAATTAATGCGG 539  
DB 504 CUAGUGCCAGACCGCGGUAUAUCGUAAGGUGUACAAGCUUAUUCGGAUAUACUGGGCG 563  
QY 540 TAAAGGTGCGCAGCGCGCTTTGTAAGTCAAGTGAATATCCCGGGCTTAACCTGGAA 599  
DB 564 UAAAGCGUGCCAGCGCGGUGUUAAGCAGUGGUGAAUUVCCCGGCGUCAAUCUGGGA 623  
QY 600 TTGCGTTGAAATCAACAAAGCTAGAGTGTGCGACAGAGGAGGTAATTCATGTGTGCA 659  
DB 624 CUGCAUUGUACUGCAAGCGCUAAGUGGCGACAGAGGAGUAGAAUUVCCCGGUGUACA 683  
QY 660 GTGAAATGCTAGAGATATGGAAGAACATCGATGGCGAAGCAGCTCTCTGGGTTAAC 719  
DB 684 GUGAAUUCGUAAGAUUUCGAGAACACCGAUGCCAGAGGCAUUVCCCGGCGUAC 743  
QY 720 TGAAGCTTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATGTCACGC 779  
DB 744 UGACGCUACAGCAAGAAAGCGUGGGAGCAAAACAGAUUAAGUACCCUGUAGUCACAGC 803  
QY 780 CCTAAACGATGCTCAATGATGTTGGGCTTATTAAGCTTGTAAGCGAAGCTTAACGCGTG 839  
DB 804 CUUAAACAUUUCACUUGUGUGUGUCUUAUCUAGUCUAGUACGUAACGAAAGCUAACGCGUG 863  
QY 840 AAGTTGACCGCGCTGGGAGATACGCTCGCAAGATTAAATCTAAAGAAATTGACGGGAGC 899  
DB 864 AAGUUGACCGGCTUGGGAGUAUCGGCCGCAAGUUGAAACUCAAAGAAUUAUGACGGGAGC 923  
QY 900 CGCACACCGGTGATTAATGATTAATGATGCAACCGGAAACCTTAATCTAATCCT 959  
DB 924 CGCAACAGCGGUGAGUAGUGUGUUAUUVGAGUACACGGGAAACCUAACCCACCU 983  
QY 960 TGACATGAGGAATTTCTAGAGTAGTATGATGCT---TCGGGAACGTTACACAGGT 1016  
DB 984 UGACAGGCGAGAACUUAUCGAGAGUGUGUGUGUGGUAAGAAAGAACCTUCACACAGGU 1043  
QY 1017 GCTGCATGCTGTGCTCAGCTCGTGTCTGATGATGTTGGTAAAGTCCCGGACGAGCGC 1076  
DB 1044 GUGUAGUGGUGUGUCUACGUCUGUGUGAGUAGUGUGUUAAGUCCCGACGAGCGC 1103  
QY 1077 AACCTTGTCAATTAATTCATCAATTGTTGGGCACTTTATGAGACTCCCGGTGACA 1136  
DB 1104 AACCCUCCAUUAGUUCUA-CAUUCAGUAGACACUUAUAGGACUGCCGGUGACAA 1162  
QY 1137 ACCGAGGAAGTGGGAGTAGACGTCAAGTCTCAATGCGCCCTAATGGGTAGGGGCTTCA 1196  
DB 1163 ACCGAGGAAGGUGGAGUAGCGUACAGUCCUACUAGGCCUUAUAGGUGGAGCUACACAC 1222  
QY 1197 GTAATACATGAGCGGCTACAGAGGTTGCCAACCCGCGAGGGAGCTTAATCTGAGAA 1256  
DB 1223 GUCAUACAUUGGUGUACAAAGGUGUCCAAACCGGAGGGGAGGCUUAUCCCAUAAAG 1282  
QY 1257 CGCGTGTATGTCGGATGCGAGTCTGCAATCTGCACTCCGTAAAGTCCGAATCGTAAGTA 1316  
DB 1283 CCAGUCGUAUUCGCGAGUCGAGUCGCAACUCGACUGGUAAGUCGGAUCCGUAAGTA 1342  
QY 1317 TCGCGAATCAACATGCTCGCGGTAATAGCTTCCGGGGCTTGTATCAACCGCCCGTCA 1376  
DB 1343 UCGUGAUCAGAAUUCACGUGAAUUAUGUUCGCGGUCUUGUACACCGCCGUGUACA 1402  
QY 1377 CCATGGAGTGGGTTTCAACGAAAGCAGTGTATACCTTAAGAGAGGCGCTTGCCACG 1436  
DB 1403 CCAUAGGAGCGGUGUUCGCGAAGUAGUAGCUUAACGUAAGGAGGCGCUUACACAG 1462  
QY 1437 GTGAGATTCAATGACTGGGGTG 1457  
DB 1463 GCGGGGUGUCGAGUCUGGGUG 1483





```
QY 1388 GGTTCACCAAGACAGAGTCTAACCTTAGAGAGGCGCTTGCCACGGTGAAT 1443
DB 1385 GGTTCACCAAGACAGAGTCTAACCTTAGAGAGGCGCTTGCCACGGTGAAT 1440

RESULT 63
ABX10819
ID ABX10819 standard; DNA; 1451 BP.
AC ABX10819;
XX
XX 17-APR-2003 (first entry)
XX
XX Burkholderia sp 16S ribosomal RNA gene.
DE Burkholderia sp 16S ribosomal RNA gene.
XX
XX Ribosomal RNA; rRNA; 16S; terephthalic acid; TA; proteobacteria;
XX TA biosynthetic pathway; isophthalic acid; 4-carboxybenzyl alcohol;
XX 4-carboxybenzaldehyde; fibre; film; paint; adhesive; beverage container;
XX gene; ds.
XX
XX Burkholderia sp.
XX
XX US6461840-B1.
XX
XX 08-OCT-2002.
XX
XX 17-JUL-2000; 2000US-00617854.
XX
XX 02-JUL-1998; 98US-0091645P.
XX PR 29-JUN-1999; 99US-00342579.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Bramucci MG, McCutchen CM, Nagarajan V, Thomas SM;
XX WPI; 2003-208774/20.
XX
XX Terephthalic acid producing proteobacteria isolated by culturing sample
XX suspected of containing the bacteria in medium having aromatic organic
XX substrate, adding xylene and isolating bacteria producing terephthalic
XX acid.
XX
XX Claim 4; Col 19-20; 21pp; English.
XX
XX The invention describes a pure isolate of a terephthalic acid (TA)
XX producing proteobacteria, comprising genes encoding all enzymes of the TA
XX biosynthetic pathway. The proteobacteria are useful for producing TA and
XX isophthalic acid, and various intermediates in the synthesis of TA such
XX as 4-carboxybenzyl alcohol and 4-carboxybenzaldehyde. The TA and
XX isophthalic acid have utility in the production of polyesters needed in
XX fibres, films, paints, adhesives and beverage containers. This sequence
XX represents the Burkholderia sp 16S ribosomal RNA gene amplification of
XX which was used to identify Burkholderia and Pseudomonas isolates
XX
XX Sequence 1451 BP; 379 A; 339 C; 441 G; 292 T; 0 U; 0 Other;
SQ
Query Match 76.0%; Score 1108; DB 10; Length 1451;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 1552; Conservative 0; Mismatches 175; Indels 9; Gaps 3;
```

```
QY 191 TTTGAGCGGCCGATGCTGATTAGCTAGTGTGGGGTAAAGCCTACCAAGCGACG 250
DB 191 TACAAAACGACCGATGTCATTAATTAATTTGGTGGGGTAAAGCTCACCAAGCGACG 250
QY 251 ATCAGTAGTGGTCTGAGAGACGACCAACCACTGGAGACTGAGACAGGCGCCAGATC 310
DB 251 ATCTGACTGTGTCGAGAGACCAACCACTGGAGACTGAGAAACGCGCCCAATCTC 310
QY 311 CTACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370
DB 311 CTACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370
QY 371 GCGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 430
DB 371 GCGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 430
QY 431 GTTAATAATCGTGAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 489
DB 431 GTTAATAATCGTGAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 489
QY 491 CAGCGCGGTAAATACGTAAGGTGCAAGGTAAATCGGAATTAATCGGAGTAAAGGTGCG 550
DB 491 CAGCGCGGTAAATACGTAAGGTGCAAGGTAAATCGGAATTAATCGGAGTAAAGGTGCG 550
QY 551 CAGCGCGGTAAATACGTAAGGTGCAAGGTAAATCGGAGTAAATCGGAGTAAATCGGAG 610
DB 551 CAGCGCGGTAAATACGTAAGGTGCAAGGTAAATCGGAGTAAATCGGAGTAAATCGGAG 610
QY 611 ACTCAAGGCTAAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 670
DB 611 ACTCAAGGCTAAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 670
QY 671 AGAATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 730
DB 671 AGAATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 730
QY 731 CACGAAAGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 790
DB 731 CACGAAAGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 790
QY 791 TCAACTAGTGTGGGCTTAAATGAGCTGTGTAACGAAAGTAAACGCGTGAAGTGAACCG 850
DB 791 TCAACTAGTGTGGGCTTAAATGAGCTGTGTAACGAAAGTAAACGCGTGAAGTGAACCG 850
QY 851 CTGGGAGAGTACGTCGCAAGATTAAACTCAAGAGAAATGACGGGAGACCGCACAGCGG 910
DB 851 CTGGGAGAGTACGTCGCAAGATTAAACTCAAGAGAAATGACGGGAGACCGCACAGCGG 910
QY 911 TGGATTATGAGATTAATTCGATGCAACGCGAAATCTTACCTTACCTTGAACATGTAGC 970
DB 911 TGGATTATGAGATTAATTCGATGCAACGCGAAATCTTACCTTACCTTGAACATGTAGC 970
QY 971 GAAATTTCTAGAGATGATTAATGCTCT--TCGGAGAGCTTAACAGAGTGTGATGCTCT 1027
DB 971 GAAATTTCTAGAGATGATTAATGCTCT--TCGGAGAGCTTAACAGAGTGTGATGCTCT 1027
QY 1028 GTGCGTACGCTGTGTCGCAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1087
DB 1028 GTGCGTACGCTGTGTCGCAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1087
QY 1088 TTAATGCTCATCATTTGGTGGGCACTTAAATGAGACTGCGGTGACCAACCGAGAGAG 1147
DB 1088 TTAATGCTCATCATTTGGTGGGCACTTAAATGAGACTGCGGTGACCAACCGAGAGAGAG 1147
QY 1148 GTGGGAGTACGTCAGTCAAGTCTCATAGGCTTAAAGGCTTACACGTAATCAATG 1207
DB 1148 GTGGGAGTACGTCAGTCAAGTCTCATAGGCTTAAAGGCTTACACGTAATCAATG 1207
QY 1208 GCGGCTACAGAGGCTGCGCAACCGCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267
DB 1208 GCGGCTACAGAGGCTGCGCAACCGCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267
QY 1268 CCGGATCGGAGTCTGCAACTGCACTCGGTGAAGTCTGGAATGCTAGTAATGCGGATCAG 1327
DB 1268 CCGGATCGGAGTCTGCAACTGCACTCGGTGAAGTCTGGAATGCTAGTAATGCGGATCAG 1327
```



|||||  
Db CCGATCGCAGTCTGCACTCGACTGGAAGCTGGAATCCCTAGTAATCCCGATCAG 1324  
Qy 1328 CATGTCGCGGTGAATACGTTCCCGGCTTTGTACACACGCCCGTCACACCATGGAGTG 1387  
Db 1325 CATGTCGCGGTGAATACGTTCCCGGCTTTGTACACACGCCCGTCACACCATGGAGTG 1384  
Qy 1388 GGTTCACCGAAGCAGTAGTCTTAACGTTAAGAGGCGCTTGCCACGGTGAGAT 1443  
Db 1385 GGTTCACCGAAGTGTAGTCTTAACGTTAAGAGGCGCTTGCCACCGGAGAT 1440

RESULT 64  
ADH74483  
ID ADH74483 standard; DNA; 1451 BP.  
XX ADH74483;  
AC ADH74483;  
XX 25-MAR-2004 (first entry)  
XX TPA producing microorganism isolate IR3 16s rRNA gene.  
XX  
XX Terephthalic acid; TPA producing microorganism; p-xylene; p-toluic acid;  
XX isophthalic acid; IPA; polyester; fibre; film; paint; adhesive;  
XX beverage container; condensation reaction; diamine; polyamide; 16s rRNA;  
XX IR3; ds.  
OS Burkholderia sp.  
XX  
XX US2003170836-A1.  
XX 11-SEP-2003.  
XX 15-AUG-2002; 2002US-00219549.  
XX 02-JUL-1998; 98US-0091645P.  
XX PR 29-JUN-1999; 99US-00342579.  
XX PR 17-JUL-2000; 2000US-00617854.  
XX (BRAM/) BRAMUCCI M G.  
XX (MCCU/) MCCUTCHEN C M.  
XX (NAGA/) NAGARAJAN V.  
XX (THOM/) THOMAS S M.  
XX Bramucci MG, McCutchen CM, Nagarajan V, Thomas SM;  
XX WPI; 2004-096737/10.  
XX  
XX New terephthalic acid-producing microorganism e.g. for producing  
XX terephthalic acid for production of polyesters commercially required in  
XX large quantities for fibers, films, paints, adhesives and beverage  
XX containers.  
XX  
XX Claim 6; SEQ ID NO 1; 24p; English.  
XX  
XX The invention relates to a terephthalic acid (TPA) producing  
XX microorganism isolated by a process comprising culturing a sample  
XX suspected of containing a TPA producing microorganism in a suitable  
XX growth medium containing at least one aromatic organic substrate selected  
XX from p-xylene, p-toluic acid and TPA, selecting those microorganisms that  
XX are able to use every substrate selected individually from p-xylene, p-  
XX toluic acid and TPA as a sole carbon source, contacting the  
XX microorganisms selected in the above step with p-xylene to form a  
XX reaction medium and monitoring the reaction medium over time for the  
XX presence of TPA, where the production of TPA indicates the presence of a  
XX TPA producing microorganism. The invention also relates to a method of  
XX producing isophthalic acid (IPA) by contacting an isolated TPA producing  
XX microorganism with an aromatic organic substrate and at least one  
XX suitable induction compound, where isophthalic acid accumulates, and  
XX optionally recovering the isophthalic acid. The microorganism is useful  
XX for producing terephthalic acid for the production of polyesters that are  
XX commercially required in large quantities for fibres, films, paints,  
XX adhesives and beverage containers. The isophthalic acid is useful in

CC condensation reactions with diamines to form polyamides. This sequence  
CC represents the 16s rRNA gene of the TPA producing microorganism  
CC Burkholderia isolate IR3.  
XX  
SQ Sequence 1451 BP; 379 A; 339 C; 441 G; 292 T; 0 U; 0 Other;  
Query Match 76.0%; Score 1108; DB 12; Length 1451;  
Best Local Similarity 87.2%; Pred. No. 0;  
Matches 1252; Conservative 0; Mismatches 175; Indels 9; Gaps 3;  
Qy 11 GCGCGCATGCTTTACACATGCAAGTCGAACGCGACACAGATGCTTGACATCTGGCGCA 70  
Db 11 GCGCGCATGCTTTACACATGCAATTCGAACGCGACACAGATGCTTGACATCTGGCGCA 70  
Qy 71 GTGGCGGACGCGGTGATTAATGATCGAAAGCTATCCAGAAAGGGGGGTAAACGATCGAA 130  
Db 71 TTGGCGAAGCGGTGATTAATGATCGAAATGATCGAAAGGGGGGTAAATCTGGCGAA 130  
Qy 131 AGATGCTTAATACCGCATATATCTTAAAGAGAAAGCAAGGAGATCGAAAGACTTGGCGC 190  
Db 131 ACCCGGATTAATACCGCATACGCTCTGAGAGAGAAAGCGGGGACCTTCGGGCTTCGCGC 190  
Qy 191 TTTTGGAGCGCGCATGATCTGATTAAGTGTGGGTAAAGGCTTAACCAAGCGAGC 250  
Db 191 TACAAAGCAGCGCGATGCAATTAATCTTGTGGGTAAAGCTCAACCAAGCGAGC 250  
Qy 251 ATCAGTAGTGTGCTGAGAGACGACCAACCACTGGGACTGAGACAGCGCCAGACTC 310  
Db 251 ATCTGACTGCTGAGAGAGACCAACCACTGGGACTGAGAAACAGCGCCCAACTC 310  
Qy 311 CTACGGGAGCAGCAGTGGGAAATTTTGGACAATGGGCCCAAGCCCTGATCCAGCAATGCC 370  
Db 311 CTACGGGAGCAGCAGTGGGAAATTTTGGACAATGGGCCCAAGCCCTGATCCCAATGCC 370  
Qy 371 GGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTCAAGAAAGGTTAGC 430  
Db 371 GGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTCAAGAAAGGTTAGC 430  
Qy 431 GTAAATATATGATGATCATGATGACGTTATCGACAGAAAGACACCGGCTTAATCTGCGCAG 490  
Db 431 GTTAATATATGATGATGATGACGTTATCGGAAATTAACACCGGCTTAATCTGCGCAG 489  
Qy 491 CAGCCCGGTTATATCTGAGGTCGAACGCTTAATCGGAATTCCTGGGCTTAAGGGTGG 550  
Db 491 CAGCCCGGTTATATCTGAGGTCGAACGCTTAATCGGAATTCCTGGGCTTAAGGGTGG 549  
Qy 551 CAGCGGCTTTGTAATGATGATGTAATCCCGGCTTAACCTGGGAAATTCGTTTGA 610  
Db 551 CAGCGGCTTTGTAATGATGATGTAATCCCGGCTTAACCTGGGAAATTCGTTTGA 609  
Qy 611 ACTACAAAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTAGCAGTGAATTCGT 670  
Db 611 ACTACAAAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTAGCAGTGAATTCGT 669  
Qy 671 AGAGATATGAGAAACATGATGCGCAAGGCAAGCTTCCTGGTTAACCTGAGCTTCAG 730  
Db 671 AGAGATATGAGAAACATGATGCGCAAGGCAAGCTTCCTGGTTAACCTGAGCTTCAG 729  
Qy 731 CAGGAAAGCTGGGAGCAAGCAAGATTGATACCTCGTAGTCCAGCGCTTAACGATG 790  
Db 731 CAGGAAAGCTGGGAGCAAGCAAGATTGATACCTCGTAGTCCAGCGCTTAACGATG 789  
Qy 791 TCAACTAGTGTGGGCTTATTAAGCTTGGTGAACGAGCTTAACGCGTGAAGTTGACGC 850  
Db 791 TCAACTAGTGTGGGCTTATTAAGCTTGGTGAACGAGCTTAACGCGTGAAGTTGACGC 849  
Qy 851 CTGGGAGTAGCGTTCGCAAGATTAAATCTCAAGAAATTGACGGGACCCGCAAGCGG 910  
Db 851 CTGGGAGTAGCGTTCGCAAGATTAAATCTCAAGAAATTGACGGGACCCGCAAGCGG 909  
Qy 911 TGATATATGATTAATTCGATGCAACGGAAGAAACCTTACCTACCTTGAATGATGAC 970  
Db 911 TGATATATGATTAATTCGATGCAACGGAAGAAACCTTACCTACCTTGAATGATGAC 969

QY 971 GAATTTCTAGAGATTAGTCT---TCGGAAACGCTTAACACAGGTCTGCATGCT 1027  
DB 970 GAATCTTGTGAGAGTGAAGTGTCTGAAAGAAACCTTAACACAGGTCTGCATGCT 1029  
QY 1028 GTCCGACGCTGCTGAGAGATGTTAGGTAAAGTCCGCAACGAGCCGCAACCTTTGCA 1087  
DB 1030 GTCCGACGCTGCTGAGAGATGTTAGGTAAAGTCCGCAACGAGCCGCAACCTTTGCTC 1089  
QY 1088 TTAATTTGCATATTGTTGGGCACTTTATATGAGACTGCGGTGACAAACCGAGAGAG 1147  
DB 1090 TTAGTTGC-----TAGCAAGAGCACTTAAGAGAGCTGCGGTGACAAACCGAGAGAG 1144  
QY 1148 GTGGAGATGACGTAAGTCTTATGAGCCCTTAATGGTAGGGCTTACACGTAATACATG 1207  
DB 1145 GTGGAGATGACGTAAGTCTTATGAGCCCTTAATGGTAGGGCTTACACGTAATACATG 1204  
QY 1208 GGGCGTACAGAGGTGGCCACCGGAGAGGGAGCTATCTCAGAAAGCGGTGATG 1267  
DB 1205 GTCCGTAACAGAGGCTGCGCAACCGGAGGTGAGCTAACCCAGAAACCGATGTAGT 1264  
QY 1268 CCGGATCGGAGTCTGCACTCGACTCCGTGAAGTCGGAATCGTATATCGCGATCAG 1327  
DB 1265 CCGGATCGGAGTCTGCACTCGACTCGGTGAAGTCGGAATCGTATATCGCGATCAG 1324  
QY 1328 CATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCACACCATGGAGTG 1387  
DB 1325 CATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCACACCATGGAGTG 1384  
QY 1386 GGTTCACCAAGAGAGGTAGTCTAACCGTAAGAGAGGGGCTTGCCACGGTGAAT 1443  
DB 1385 GGTTCACCAAGAGGTAGTACCTTAACCGTAAGAGAGGGGCTTGCCACGGCAGAT 1440

## RESULT 65

ABL53377  
ID ABL53377 standard; DNA; 1453 BP.

XX ABL53377;

XX 27-JUN-2002 (first entry)

DE Escherichia coli 16S DNA sequence #2.

KM Arsenic oxidation; 16S; water treatment; ds.

XX Escherichia coli.

PN FR2814751-A1.

PD 05-APR-2002.

PF 03-OCT-2000; 2000FR-00012579.

PR 03-OCT-2000; 2000FR-00012579.

XX (REGE-) BUREAU RECH GEOLOGIQUES & MINE.

PI Battaglia BF, Dictor WC, Morin D, Baranger P;

DR WPI; 2002-332477/37.

PT New bacteria that oxidize arsenic(III), useful for treating drinking  
water or effluents, can use arsenic(III) as sole energy source.

PS Claim 1; Page 27; 28pp; French.

CC The present invention relates to isolated, autotrophic, aerobic bacteria  
that can oxidize arsenic (III) to arsenic (V) using carbon dioxide as its  
sole source of carbon and as (III) as sole source of energy. They have  
16S sequences (ABL53376 and ABL53377). The bacteria are useful for  
treating water intended for human consumption, or mining and industrial  
waste, and for converting arsenic (III) to arsenic (V) which is easier to

CC remove, less toxic and less mobile

XX Sequence 1453 BP; 372 A; 340 C; 465 G; 276 T; 0 U; 0 Other;

Query Match 75.8%; Score 1104.4; DB 6; Length 1453;

Best Local Similarity 86.2%; Pred. No. 0; Mismatches 191; Indels 10; Gaps 3;

QY 1 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCCGACGACGATGCTTGAT 60  
DB 1 ATTGAACGCTGGCGGATGCTTTACACATGCAAGTCCGACGACGATGCTTGAT 58  
QY 61 CTGGTGGAGATGGGCGGAGGAGTGAATGATCATCGAACGTAATCCAGAAAGAGGGAGTA 120  
DB 59 CTGGTGGAGATGGGCGGAGGAGTGAATGATCATCGAACGTAATCCAGAAAGAGGGAGTA 118  
QY 121 ACGCATCGAAGATGCTTAATACCGCATATATCTTAAGAGAAAGCAGGGAGTCAAA 180  
DB 119 ACTGGCGAAGACGCTGCTTAATACCGCATATATCTTAAGAGAAAGCAGGGAGTCAAA 178  
QY 181 GACCTTGCGCTTTTGGAGCGGCCGATGTTCTGATTAAGTGTGGGTAAAGGCTTAC 240  
DB 179 GACCTTGCGCTTTTGGAGCGGCCGATGTTCTGATTAAGTGTGGGTAAAGGCTTAC 238  
QY 241 CAAGCGCAGATCGATGTTGTTGCTGAGAGAGACGACGACGACGACGACGACGACGACG 300  
DB 239 CAAGCGCAGATCGATGTTGTTGCTGAGAGAGACGACGACGACGACGACGACGACGACG 298  
QY 301 GCCCAGACTCTTACGAGAGGAGCAGCAGTGGGAAATTTTGAACATGGCGCAAGCTGATC 360  
DB 299 GCCCAGACTCTTACGAGAGGAGCAGCAGTGGGAAATTTTGAACATGGCGCAAGCTGATC 358  
QY 361 CAGCAATGCCCGTGAAGTGAAGAAAGGCTTGGGTGTTAAAGCTCTTCACTCGAAGAA 420  
DB 359 CAGCAATGCCCGTGAAGTGAAGAAAGGCTTGGGTGTTAAAGCTCTTCACTCGAAGAA 418  
QY 421 AAAGTTACGGTAAATATGTTGATCTGATGATGATGATGATGATGATGATGATGATGAT 480  
DB 419 AATGCTGCTGCTTAATACACGATGATGATGATGATGATGATGATGATGATGATGATGAT 478  
QY 481 TACGTCGACGAGCGCGGTAATACGTAAGGTGCAAGGCTTAATCGAATTAAGTGGCGGT 540  
DB 479 TACGTCGACGAGCGCGGTAATACGTAAGGTGCAAGGCTTAATCGAATTAAGTGGCGGT 538  
QY 541 AAAGGTGCGCAGGCGCTTGTAAAGTCAATGTGAATCCCGGGCTTAACTCGGGAAT 600  
DB 539 AAAGGTGCGCAGGCGCTTGTAAAGTCAATGTGAATCCCGGGCTTAACTCGGGAAT 598  
QY 601 TGCGTTGAATCTAAGGCTAGAGGTGTCAGAGAGGAGGTGAATTCATGTGTAGCAG 660  
DB 599 TGCGTTGAATCTAAGGCTAGAGGTGTCAGAGAGGAGGTGAATTCATGTGTAGCAG 658  
QY 661 TGAATGCTGAGATATGAAAGAACTCATGTGCGAAGGCGCTCTGGGTTAACT 720  
DB 659 TGAATGCTGAGATATGAAAGAACTCATGTGCGAAGGCGCTCTGGGTTAACT 718  
QY 721 GACGCTCATGCAAGAAAGCTGGGAGCAACAGATTAATACCTCGTGTGTGACGCC 780  
DB 719 GACGCTCATGCAAGAAAGCTGGGAGCAACAGATTAATACCTCGTGTGTGACGCC 778  
QY 781 CTAAACGATGTCATGTTGTTGGGCTTTAAGGCTGGTAAAGAACTTAAGCGCGTGA 840  
DB 779 CTAAACGATGTCATGTTGTTGGGCTTTAAGGCTGGTAAAGAACTTAAGCGCGTGA 838  
QY 841 AGTTGACCGCTGGGAGTACGCTGCAAGATTAATACTCAAGAAATTGAACGGGAGACC 900  
DB 839 AGTTGACCGCTGGGAGTACGCTGCAAGATTAATACTCAAGAAATTGAACGGGAGACC 898  
QY 901 GCACAAAGGATGATTAATGATTAATTCATGATCAACGCAAAACCTTAACCTT 960  
DB 899 GCACAAAGGATGATTAATGATTAATTCATGATCAACGCAAAACCTTAACCTT 958  
QY 961 GACATGTAAGCAATTTTCTAAGATTAATGATGCT---TCGGGAACGCTTAACAGGTG 1017

```

Db      959  GACATGCTGGAATCCCGAAGATTGGAGTGTCCGAAAGAGCCAGAACAGAGTG 1018
Qy      1018  CTGCATGGCTGCTGTCAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAAGAGCGCA 1077
Db      1019  CTGCATGGCTGCTGTCAGCTCGTGTGATGATGTTGGTTAAGTCCCGCAAGAGCGCA 1078
Qy      1078  ACCCTGTGATTAAATGTCATCATATTGGTGGCACTTTAATGAGACTGCCGTGACAA 1137
Db      1079  ACCCTGTGATTAAATGTCATCATATTGGTGGCACTTTAATGAGACTGCCGTGACAA 1133
Qy      1138  CCGGAGGAAGTGGGGATACGTCAAGTCCATGAGCCCTTATGAGTGGCTTACACAG 1197
Db      1134  CCGGAGGAAGTGGGGATACGTCAAGTCCATGAGCCCTTATGAGTGGCTTACACAG 1193
Qy      1198  TAATCAATGAGCGCGATGAGAGGTTGCCAACCCCGGAGGGAGCTAATCTCAGAAAC 1257
Db      1194  TAATCAATGAGCGCGATGAGAGGTTGCCAACCCCGGAGGGAGCCATTCCTAATAAC 1253
Qy      1258  GCGTGTAGTCCGATCGAGTCTGCAACTGCACTCGTGAAGTGGAAATCGTATGAT 1317
Db      1254  CGTGTAGTCCGATCGAGTCTGCAACTGCACTCGTGAAGTGGAAATCGTATGAT 1313
Qy      1318  CCGGATCAGCATGTGCGGTGATACGTTCCCGGGTCTTGTACACACCGCCCGTCAAC 1377
Db      1314  CCGGATCAGCATGTGCGGTGATACGTTCCCGGGTCTTGTACACACCGCCCGTCAAC 1373
Qy      1378  CATGGAGAGGGGTTTACAGGAAGCATAGTCTAACCGTAAAGAGGCGCTTGCCACG 1437
Db      1374  CATGGAGAGGGGTTTACAGGAAGCATAGTCTAACCGTAAAGAGGCGCATTAACCG 1433
Qy      1438  TGAGATTCATGACTGGGGTG 1457
Db      1434  TAGGCTTCATGACTGGGGTG 1453

RESULT 66
ADG44143/c
ID      ADG44143 standard; DNA; 1453 BP.
AC      ADG44143;
AC      26-FEB-2004 (first entry)
DE      Unknown organism strain RA2 16S rDNA.
KW      de; rDNA; bacterial strain identification; wastewater bioreactor; 16S.
OS      unidentified.
XX      US2003207321-A1.
XX      06-NOV-2003.
XX      17-JUN-2003; 2003US-00464709.
XX      PF      16-DEC-1999; 99US-0171140P.
XX      PR      13-DEC-2000; 2000US-00735567.
XX      PA      (BRAM/) BRAMUCCI M G.
XX      PA      (CHEN/) CHEN M W.
XX      PA      (ALBE/) ALBERTSON-KANE H M.
XX      PA      (NAGM/) NAGARAJAN V.
XX      PI      Bramucci MG, Chen MW, Albertson-Kane HM, Nagarajan V;
XX      DR      WPI; 2003-901056/82.
XX      PT      New isolated 16S rDNA sequence indicative of the presence of an activated
XX      sludge bacterial strain, useful for rapidly identifying and monitoring
XX      bacterial strains in samples taken from a wastewater bioreactor.
XX      Claim 1; SEQ ID NO 4; 24bp; English.

```

```

XX      CC      The invention relates to an isolated 16S rDNA sequence. The 16S rDNA
CC      sequences and methods are useful in rapidly identifying and monitoring
CC      the corresponding bacterial strains in samples taken from a wastewater
CC      bioreactor or from any environment which might contain the specified
CC      bacteria. The present sequence represents the 16S rDNA from an unknown
CC      organism.
SQ      Sequence 1453 BP; 279 A; 469 C; 344 G; 361 T; 0 U; 0 Other;
Query Match      75.7%; Score 1103.2; DB 10; Length 1453;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 188; Indels 12; Gaps 3;
Qy      1      ATTGAACGCTGGCGGCATGCTTTTACACATGCAAGTGCAGACGGCAGCAGGATGCTTGAT 60
Db      1449  ATTGAACGCTGGCGGCATGCTTTTACACATGCAAGTGCAGACGGCAGCAGGATGCTTGAT 1394
Qy      61      CTGGTGGGAGTGGGCGGAGCGGGTGAATGATCGAAGCATGACAGAGAGGGGGTA 120
Db      1393  CTGGCGGAGAGCGGAGCGGGTGAATGATCGAAGCATGACAGAGAGGGGGTA 1334
Qy      121      ACGCATCGAAGATGTGCTAATACCGCATATATCTTAAAGAGAAAGAGGAGATCGAA 180
Db      1333  GCCCGCGAAGCGGATTAATACCGCATGTGATCGAAGTGAAGTGGGGAGCCGCA 1274
Qy      181      GACCTTGGCTTTTGGAGCGCGCATGTCTGATTAGTATTGTTGGTGGGTAAAGGCTTAC 240
Db      1273  GGCTTCAGCGCTTTGAGCGCGCATGTGATGATTGTTGGTGGGTAAAGGCTTAC 1214
Qy      241      CAAGCGAGCATCACTAATGTTGCTGAGAGAGCGACGACCACTGGGAGCTGAGACAG 300
Db      1213  CAAGCGAGCATCACTAATGTTGCTGAGAGAGTACGACCACTGGGAGCTGAGACAG 1154
Qy      301      GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGATTTTGGACAAATGGGCGCAAGCTGATC 360
Db      1153  GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGATTTTGGACAAATGGGCGCAAGCTGATC 1094
Qy      361      CAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGTAAAGCTTTTCACTGAGAGAGA 420
Db      1093  CAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGTAAAGCTTTTGTGAGAGAGA 1034
Qy      421      AAGGTTACGGTAATTAATCGTACTCATAGAGGTATGACAGAGAGAGACCGGCTAAC 480
Db      1033  AAGGTTACGGTAATTAATCGTACTCATAGAGGTATGACAGAGAGAGACCGGCTAAC 974
Qy      481      TAGTGCACGAGCGCGGTAAATAGTGAAGGTGCAAGCGTTATCGAATTACTGGGCGT 540
Db      973  TAGTGCACGAGCGCGGTAAATAGTGAAGGTGCAAGCGTTATCGAATTACTGGGCGT 914
Qy      541      AAGGTTGCGCAGCGCGCTTTGTAAATGATGATGAAATCCCGGGCTTAACTTGGAAAT 600
Db      913  AAGGTTGCGCAGCGCGCTTTGTAAATGATGATGAAATCCCGGGCTTAACTTGGAAAC 854
Qy      601      TCGCTTTGAACTACAGAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTATGACG 660
Db      853  TCGCTTTGAACTACAGAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTATGACG 794
Qy      661      TGAATGCGTAGATGATGAGAAACATGATGCGAAGGAGCGCTCTGGGTAACTACT 720
Db      793  TGAATGCGTAGATGATGAGAAACATGATGCGAAGGAGCGCTCTGGGTAACTACT 734
Qy      721      GACGTCATGACGAAAGCGTGGGAGAGCAAGAGATTAGATACCTGTAGTCCAGCGC 780
Db      733  GACGTCATGACGAAAGCGTGGGAGAGCAAGAGATTAGATACCTGTAGTCCAGCGC 674
Qy      781      CTAACAGATGATCACTAGTTGTTGGGCTTATAGCTTGTGTAAGAACTAACGGGTGA 840
Db      673  CTAACAGATGATCACTAGTTGTTGGGCTTATAGCTTGTGTAAGAACTAACGGGTGA 614
Qy      841      AGTTGACCGCTGGGGAGTACGGTGCAGAGTTAAACTCAAGAGATTGACGGGAGCCC 900
Db      613  AGTTGACCGCTGGGGAGTACGGTGCAGAGTTAAACTCAAGAGATTGACGGGAGCCC 554

```



Db 1333 GCCCGCGAAGCCCGATTATATACCGCATGTGATCTGAGATGAAAGTGGGAGCCGCA 1274  
 Qy 181 GACCTTGGCTTTTGGAGCGCGCATATGTCTGATTAGCTAGTGTGGGCTAAAGCCTAC 240  
 Db 1273 GGCCTCAGCGCTTTGGAGCGCGCATATGTAGTGTGGTGGGCTAAAGCCTTAC 1214  
 Qy 241 CAAGCGCATGATCAGTATGTGTCTGAGAGACGACACCGCACTCTGGGACTGAGACAG 300  
 Db 1213 CAAGCTGCGATCTGTAGCTGTGTGAGAGATGATCAGCCACTCTGGGACTGAGACAG 1154  
 Qy 301 GCCCAGACTCTTACCGGAGGCGACGATGGGGAATTTTGGACATATGGGCCCAAGCCTTATC 360  
 Db 1153 GCCCAGACTCTTACCGGAGGCGACGATGGGGAATTTTGGACATATGGGCCCAAGCCTTATC 1094  
 Qy 361 CAGCATATGCGCGGTGATGATGAAGGCGCTTGGGCTTGAAGCTCTTTCAGTGAAGAGA 420  
 Db 1093 CAGCCATTCGCGGTGATGATGAAGGCGCTTGGGCTTGAAGCTCTTTCAGTGAAGAGA 1034  
 Qy 421 AAGGTTACGCTAATATATCTGTGATCATGACGCTATCGACAGAGAGACACCGGCTAAC 480  
 Db 1033 AAGGCTCTCTCTAATATACGAGGCGCATATGACGCTACCGTAAGATATAGACCGGCTAAC 974  
 Qy 481 TACGTCCAGACAGCCCGGTATATGATGAGGTGCAAGCCTTATTCGGAATTAATCTGGGCGT 540  
 Db 973 TACGTCCAGACAGCCCGGTATATGATGAGGTGCAAGCCTTATTCGGAATTAATCTGGGCGT 914  
 Qy 541 AAGGTTGCGGAGCGCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
 Db 913 AAGGCTGCGGAGCGCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 854  
 Qy 601 TGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 Db 853 TGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794  
 Qy 661 TGAATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 Db 793 TGAATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 734  
 Qy 721 GACGCTCAGACAGAAAGCGTGGGAGCAACAGATTAGATACCTGGTATGACAGCC 780  
 Db 733 GACGCTCAGACAGAAAGCGTGGGAGCAACAGATTAGATACCTGGTATGACAGCC 674  
 Qy 781 CTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 Db 673 CTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614  
 Qy 841 AGTTGACCGCTGGGAGTACGCTCGCAGATTTAACTCAAGAAATTGACCGGAGCC 900  
 Db 613 AGTTGACCGCTGGGAGTACGCTCGCAGATTTAACTCAAGAAATTGACCGGAGCC 554  
 Qy 901 GCAACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 Db 553 GCAACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494  
 Qy 961 GACATGAGGAATTTCTAGAGATAGATTAGTCT---TCGGGAGCGCTTAACAGAGT 1017  
 Db 493 GACATGAGGAATTTCTAGAGATAGATTAGTCT---TCGGGAGCGCTTAACAGAGT 434  
 Qy 1018 CTGATGAGCTGTCTGACGCTGTCTGTGTGATGATGATGATGATGATGATGATGATGATGAT 1077  
 Db 433 CTGATGAGCTGTCTGACGCTGTCTGTGTGATGATGATGATGATGATGATGATGATGATGAT 374  
 Qy 1078 ACCCTGTCTATTAATGCTCATATTTGGTGGGAGCTTTATGAGACTGCGCGGTGACAA 1137  
 Db 373 ACCCTGTCTATTAATGCTCATATTTGGTGGGAGCTTTATGAGACTGCGCGGTGACAA 319  
 Qy 1138 CCGGAGGAGGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGAGGTAGGGCTTACACG 1197  
 Db 318 CCGGAGGAGGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGAGGTAGGGCTTACACG 259  
 Qy 1198 TAATACATGCGCGGTGACAGAGGTTGGCAACCGCGAGGGGAGCTTAATTCAGAAAGC 1257  
 Db 258 TCATACATGCGCGGTGACAGAGGTTGGCAACCGCGAGGGGAGCTTAATTCAGAAAGC 199

Qy 1258 GCGTCTGATCCCGATCGGATCTTGCACTGACCTCCGTAAGTGGGAATGCTAGTAT 1317  
 Db 198 CAGTCTGATCCCGATCGGATCTTGCACTGACCTCCGTAAGTGGGAATGCTAGTAT 139  
 Qy 1318 CCGGATCAGATGCGGAGTGAATGCTTCCGGGTCTTGTACACACCGCCGTACAC 1377  
 Db 138 CCGGATCAGATGCGGAGTGAATGCTTCCGGGTCTTGTACACACCGCCGTACAC 79  
 Qy 1378 CATGGAGTGGGTTTACACAGAGCGATGCTTAACCGTAAAGAGGCGCTTGCACG 1437  
 Db 78 CATGGAGCGGTTCTGCAGAAAGTATAGCTTAACCGCAAGAGGCGGATTAACACG 19  
 Qy 1438 TGAGATTATGACTGGG 1455  
 Db 18 CAGGTTCTGACTGGG 1  
 RESULT 68  
 ADL27933/C  
 ID ADL27933 standard; DNA; 1453 BP.  
 XX  
 AC ADL27933;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE RA2 16S rDNA.  
 XX  
 KW Activated sludge bacterial strain; pyruvate; degrading bacterial strain;  
 KW acetate; benzoate; nitrate; wastewater sludge; groundwater; soil;  
 KW nitrate reduction; RA2; 16S rDNA; ds.  
 OS Unidentified.  
 XX  
 PN US2003207320-A1.  
 XX  
 PD 06-NOV-2003.  
 XX  
 PF 17-JUN-2003; 2003US-00464356.  
 XX  
 PR 16-DEC-1999; 99US-0171140P.  
 PR 13-DEC-2000; 2000US-00735567.  
 XX  
 PA (BRAM/) BRAMUCCI M G.  
 PA (CHEN/) CHEN M W.  
 PA (ALBE/) ALBERTSON-KANE H M.  
 PA (NAGA/) NAGARAJAN V.  
 XX  
 XX Bramucci MG, Chen MW, Alberson-Kane HM, Nagarajan V;  
 PI WPI; 2003-901055/82.  
 DR  
 PT New isolated 16S rDNA sequence indicative of the presence of an activated  
 PT sludge bacterial strain, useful for rapidly identifying and monitoring  
 PT bacterial strains in samples taken from a wastewater bioreactor.  
 PS  
 XX Claim 1; SEQ ID NO 4; 25bp; English.  
 CC  
 CC The invention relates to an isolated 16S rDNA sequence. The isolated 16S  
 CC rDNA sequence is indicative of the presence of an activated sludge  
 CC bacterial strain. The invention also relates to methods for identifying  
 CC an activated sludge bacterial strain, a pyruvate degrading bacterial  
 CC strain, an acetate degrading bacterial strain, a benzoate degrading  
 CC bacterial strain, methods for degrading pyruvate, acetate or benzoate and  
 CC a method for reducing nitrate. Identifying an activated sludge bacterial  
 CC strain comprises extracting genomic DNA from a bacterial cell and probing  
 CC the extracted genomic DNA with a probe derived from any of the above-  
 CC mentioned sequences under suitable hybridisation conditions, where the  
 CC identification of a hybridisable nucleic acid fragment confirms the  
 CC presence of an activated sludge bacterial strain. Alternatively, the  
 CC method comprises extracting genomic DNA from a bacterial cell and  
 CC amplifying the extracted genomic DNA with at least one oligonucleotide  
 CC primer corresponding to a portion of the above-mentioned sequences, such

CC that amplification products are generated, where the presence of  
CC amplification products confirms the presence of an activated sludge  
CC bacterial strain. The bacterial cell is contained within wastewater  
CC sludge or within an environmental sample selected from groundwater and  
CC soil. Degrading pyruvate, acetate or benzoate comprises contacting the  
CC bacterial strain under suitable growth conditions with an amount of  
CC pyruvate, acetate or benzoate. Reducing nitrate comprises contacting the  
CC bacterial strain under suitable growth conditions with an amount of  
CC nitrate. The 16S rDNA sequences and methods are useful in rapidly  
CC identifying and monitoring the corresponding bacterial strains in samples  
CC taken from a wastewater bioreactor or from any environment which might  
CC contain the specified bacteria. This sequence represents an isolated 16S  
CC rDNA of the invention.

XX Sequence 1453 BP; 279 A; 469 C; 344 G; 361 T; 0 U; 0 Other;

Query Match 75.7%; Score 1103.2; DB 11; Length 1453;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 1258; Conservative 0; Mismatches 188; Indels 12; Gaps 3;

QY 1 ATTGAACGCTGGCGGCGATCTTTACATGACATGCAAGCGACGACGATGCTTCAT 60  
DB 1449 ATTGAACGCTGGCGGCGATCTTTACATGACATGCAAGCGGCGGCGCA----AC 1394  
QY 61 CTGGGCGGCGAGTGGCGGCGGCGGATGATGCGAAGCGATCCAGAAAGAGGCGGATG 120  
DB 1393 CTGGGCGGCGAGGCGGCGGCGGATGATGCGAAGCGATCCAGAAAGAGGCGGATG 1334  
QY 121 ACGCATCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGAGGCGGATG 180  
DB 1333 GCCCGCGGAAACCGGATTAATACCGCATGTGATCTGAGGATGAAAGTGGGCGGCA 1274  
QY 181 GACCTTGGCGCTTTTGAAGCGCGCGATGTCTGATTAGCTGTTGGTGGGTTAAAGGCTTAC 240  
DB 1273 GGCCTTCAGCGCTTTTGAAGCGCGCGATGTGCGATTGCTGTTGGGTTAAAGGCTTAC 1214  
QY 241 CAAGCGAGCATGATGATGTTGCTGAGAGAGAGCGACGACCTGCGGCTGAGACAG 300  
DB 1213 CAGGCTGCGATCTGTAGCTGCTGAGAGATGATCAGCCACCTGGGCTGAGACAG 1154  
QY 301 GCCCGACTCTCTACCGGAGAGGAGCAGATGGGGAATTTTGGACAATGGGCGCAAGCTTATC 360  
DB 1153 GCCCGACTCTCTACCGGAGAGGAGCAGATGGGGAATTTTGGACAATGGGCGCAAGCTTATC 1094  
QY 361 CAGCATGCGCGGCTGATGAGAGAGGCGCTTGGGTTTGAAGCTCTTTTCACTGAGAGAG 420  
DB 1093 CAGCATTCGCGGCTGAGAGAGGCGCTTGGGTTTGAAGCTCTTTTGAAGAGAGAG 1034  
QY 421 AAAGGTTACGTAATTAATCGTACATGACGCTATGACAGAGAGAGCAAGCGCTAAC 480  
DB 1033 AAAGGCTCTCTTAATACAGGGGCGCATGACGCTACCGTAAGATTAAGACACCGGCTAAC 974  
QY 481 TACGTGCAAGACCGCGGCTTAATAGTGGGTGCAAGCGTAAATCGAATTAATCTGGGCT 540  
DB 973 TACGTGCAAGACCGCGGCTTAATAGTGGGTGCAAGCGTAAATCGAATTAATCTGGGCT 914  
QY 541 AAAGGTTGCGAGCGGCTTTTGAAGCTGATGATGTAATCCCGGCGCTTAAGCTGGAGAT 600  
DB 913 AAAGGTTGCGAGCGGCTTTTGAAGCTGATGATGTAATCCCGGCGCTTAAGCTGGAGAT 854  
QY 601 TGGCTTTGAATCTAAGGCTAGAGTGGCAGAGAGAGTGAATTCATATGTTAGAGAG 660  
DB 853 TGGCTTTGATGCTAGAGGCTGAGTGGCGGAGAGAGAGTGAATTCATATGTTAGAGAG 794  
QY 661 TGAATGCGTATGATATGAGAGAGCATGATGCGAAGGCGAGCTCTCTGGGTTAACT 720  
DB 793 TGAATGCGTATGATATGAGAGAGCATGATGCGAAGGCGAGCTCTCTGGGCTGAGACT 734  
QY 721 GACGCTCATGACGAGAAAGCTGGGAGAGAGAGATTAATGATCCCTGGTATGTCACGCG 780  
DB 733 GACGCTCATGACGAGAAAGCTGGGAGAGAGAGATTAATGATCCCTGGTATGTCACGCG 674  
QY 781 CTAAACGATGTCACTAGTTGTTGGGCTTATTAAGCTTTGTAAGAGAGCTAACCGTGA 840

DB 673 CTAAACGATGTCACTAGTTGTTGGGATTTATTTCTCACTTAAGAGAGCTAACGCGTGA 614  
QY 841 AGTTGACCGCTCTGGGAGATACGCTGCGCAAGTTAAATCTCAAGAGATTAAGCGGAGACC 900  
DB 613 AGTTGACCGCTCTGGGAGATACGCGCGCAAGGTTAAATCTCAAGAGATTAAGCGGAGACC 554  
QY 901 GCAAGAGCGGATTAATGATGATTAATGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
DB 553 GCAAGAGCGGATTAATGATGATTAATGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494  
QY 961 GACATGTACCAATTTTCTAGAGATGATGATGCT--TCGGAGAGCGCTTAACAGAGTG 1017  
DB 493 GACATGTACCAATTTTCTAGAGATGATGATGCT--TCGGAGAGCGCTTAACAGAGTG 434  
QY 1018 CTGATGAGCTGCTGCTGAGCTGCTGCTGAGATGTTGGGTTAAATCCCGCAAGAGCGCA 1077  
DB 433 CTGATGAGCTGCTGCTGAGCTGCTGCTGAGATGTTGGGTTAAATCCCGCAAGAGCGCA 374  
QY 1078 ACCCTGTCAATTAATGCGATCATTTGTTGGGCACTTAATGAGACTGCGGAGAGCAAA 1137  
DB 373 ACCCTGTCAATTAATGCGATCATTTGTTGGGCACTTAATGAGACTGCGGAGAGCAAA 319  
QY 1138 CCGAGAGAGAGTGGGAGATGAGCTCAAGTCTCATGCGCTTAATGAGAGCTTACACAG 1197  
DB 318 CCGAGAGAGAGTGGGAGATGAGCTCAAGTCTCATGCGCTTAATGAGAGCTTACACAG 259  
QY 1198 TAAATCAATGCGGCTACAGAGGCTTCCAAACCGCGAGGAGAGCTTAATTCAGAAAGC 1257  
DB 258 TCATATCAATGCTGCTGCTCAAGAGGCTTCCAAACCGCGAGGAGAGCTTAATTCAGAAAGC 199  
QY 1258 GCGTGTGATGCGGATGAGAGTCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 1317  
DB 198 CAGTGTGATGCTGCGATGCGATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 139  
QY 1318 CGCGGATCAGCATGCTGCGGATGAAATGCTTCCCGGCTTGTACACACCGGCGCTCACAC 1377  
DB 138 CGTGGATCAGCATGCTGCGGATGAAATGCTTCCCGGCTTGTACACACCGGCGCTCACAC 79  
QY 1378 CATGGAGTGGGCTTACACAGAGAGAGTACTTAATCCGTAAGAGAGGCGCTTGCACAG 1437  
DB 78 CATGGAGTGGGCTTCTGCAAGAGTAAATTAAGCTTAACCGCAAGAGAGGCGATTAACACAG 19  
QY 1438 TGAATTCATGACGCGG 1455  
DB 18 CAGGCTGTGATGCGG 1

RESULT 69  
ADP47789/C  
ID ADP47789 standard; DNA; 1453 BP.  
XX  
XX ADP47789;  
AC 12-FEB-2004 (first entry)  
XX  
XX  
DT 12-FEB-2004 (first entry)  
XX  
XX  
DE Unknown bacterial strain RA2 16S rDNA.  
XX  
XX 16S rDNA; activated sludge; bacterial strain; diagnosing organism;  
XX unrecognised bacteria; industrial wastewater bioreactor; ds.  
XX  
XX  
OS Unidentified.  
XX  
XX  
FN US6608190-B1.  
XX  
XX  
PD 19-AUG-2003.  
XX  
XX  
PE 13-DEC-2000; 2000US-00735567.  
XX  
XX 16-DEC-1999; 99US-0171140P.  
XX  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX





XX (MARTIN P A W.  
 PA (SHRO) SHROPSHIRE A D S.  
 PA (GUND) GUNDERSEN-RINDAL D E.  
 PA (BLAC) BLACKBURN M B.  
 PI Martin PAM, Shropshire ADS, Gundersen-Rindal DE, Blackburn MB;  
 XX WPI: 2005-283752/29.  
 DR GENBANK: AY344056.  
 XX  
 PT New biologically pure culture of a strain of *Chromobacterium suttsuga* sp.  
 PT nov. possessing insecticidal activity, useful for biocontrol of insect  
 PT pests.  
 PS Claim 3; SEQ ID NO 1; 19pp; English.  
 XX  
 CC The present invention relates to a biologically pure culture of a strain  
 CC of *Chromobacterium suttsuga* sp. nov. The invention also relates to  
 CC insecticidally active metabolites obtained from the strain of  
 CC *Chromobacterium suttsuga* sp. nov. The invention is useful for biocontrol  
 CC of insect pests. The present sequence is the *Chromobacterium suttsuga*  
 CC NRRL B-30655 16S ribosomal DNA (rDNA).  
 XX  
 SQ Sequence 1440 BP; 359 A; 329 C; 460 G; 291 T; 0 U; 1 Other;  
 Query Match 75.7%; Score 1103; DB 14; Length 1440;  
 Best Local Similarity 87.5%; Pred. No. 0;  
 Matches 1441; Conservative 1; Mismatches 172; Indels 5; Gaps 3;

QY 1 ATTGAACGCTGGCGGAGCTTTACACATGCAAGTGAAGCGGACAGCGATGCTTCAT 60  
 DB 24 ATTGAACGCTGGCGGAGCTTTACACATGCAAGTGAAGCGGATGCA -GGTGTCTTGAC 82  
 QY 61 CTGTGTGCGAGTGTGGCGGAGCGGCTGAGTATGATCGGAACCTATCCAGAAAGGCGGCTA 120  
 DB 83 C-GCTGAGAGTGTGGGAAAGCGGCTGAGTATGCTGTGGAATGTAACGTGTATATGGGGAAT 141  
 QY 121 ACGCATCGAAGATGTCTTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAA 180  
 DB 142 GCTCGCGGAAAGCCGATTAATACCGCATATCCCTGAGGGGAAAGCGGGGATCGAAA 201  
 QY 151 GACCTTGGCGCTTTGAGAGCGCGGATGTCTGATTAAGCTAGTGTGGGTAAGGCTTAC 240  
 DB 202 GACCTTGGCGCTTTGAGAGCGCGGATGTCTGATTAAGCTAGTGTGGGTAAGGCTTAC 261  
 QY 241 CAAGCGAGCATGATGTTGTCTGAGAGAGAGCAACAGCCACACTGGGACTGAGACAG 300  
 DB 252 CAAGCGAGCATGATGTTGTCTGAGAGAGAGTATCCGCCACACTGGGACTGAGACAG 321  
 QY 301 GCCGAGACTCTTACGGGAGGAGCAGTGTGGGAAATTTGGACAATGGGCGCAAGCTGATC 360  
 DB 322 GCCGAGACTCTTACGGGAGGAGCAGTGTGGGAAATTTGGACAATGGGCGCAAGCTGATC 381  
 QY 351 CAGCAATGCGCGGTAGTGAAGAGCCCTTGGGTTGTAAGCTCTTTCAGTCGAGAGA 420  
 DB 382 CAGCAATGCGCGGTAGTGAAGAGCCCTTGGGTTGTAAGCTCTTTCAGTCGAGAGA 441  
 QY 421 AAAGGTTACGTTAAATATCGTACTCATGACGCTATCGACAGAAAGACACCGGCTAAC 480  
 DB 442 AATCCCGCTGTATTAATACCGCGGGAGTGAAGTACBGAAGAAATTAAGCACCGGCTAAC 501  
 QY 481 TAGGTGCAAGAGCGCGGCTTAATACGTAGGGTGAAGCGTTATTCGGAATTAATCTGGCGCT 540  
 DB 502 TAGGTGCAAGAGCGCGGCTTAATACGTAGGGTGAAGCGTTATTCGGAATTAATCTGGCGCT 561  
 QY 541 AAAGGTGCGAGCGGCGCTTGTAAAGTCAAGTGAATCCCGGCGTTAACCTGGGAAT 600  
 DB 562 AAAGGTGCGAGCGGCGCTTGTAAAGTCAAGTGAATCCCGGCGTTAACCTGGGAAT 621  
 QY 601 TGCGTTTGAACCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGCAG 660  
 DB 622 GGCAATTGAGAGCTGAGAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGCAG 681

QY 661 TGAATGCGTAGAGATATGAAAGACATGATGCGAAGGAGCGGCTCTGGGTTAACACT 720  
 DB 682 TGAATGCGTAGAGATATGAAAGACATGATGCGAAGGAGCGGCTCTGGGTTAACACT 741  
 QY 721 GACGCTATGACAGAAAGCGTGGGAGAGCAACAGATTAGATACCTGTAGTCCAGCC 780  
 DB 742 GACGCTATGACAGAAAGCGTGGGAGAGCAACAGATTAGATACCTGTAGTCCAGCC 801  
 QY 781 CTAAACGATGTCAACTAGTGTGGGCTTTATAGGCTGTGTAACGAAGCTAACCGCTGA 840  
 DB 802 CTAAACGATGTCAACTAGTGTGGGCTTTATAGGCTGTGTAACGAAGCTAACCGCTGA 861  
 QY 841 AGTTGACCGCTGGGAGATAGTGTGCAAGATTTAACTCAAGAAATTTGACGGGAGCC 900  
 DB 862 AATTGACCGCTGGGAGATAGTGTGCAAGATTTAACTCAAGAAATTTGACGGGAGCC 921  
 QY 901 GCACAAAGCGGTGATTTATGATGATTTATTCGATGCAACGCGAAAAAAGCTTACCTT 960  
 DB 922 GCACAAAGCGGTGATTTATGATGATTTATTCGATGCAACGCGAAAAAAGCTTACCTT 981  
 QY 961 GACATGTAGCAATTTTCTAAGATATGATGCTGCTTGGGAAAGCTTAAACAGGTG 1017  
 DB 982 GACATGTAGCAATTTTCTAAGATATGATGCTGCTTGGGAAAGCTTAAACAGGTG 1041  
 QY 1018 CTGATGCGTGTGCTGAGCTGTGCTGATGATGATGCTGCTTAAAGTCCGCAAGAGCGCA 1077  
 DB 1042 CTGATGCGTGTGCTGAGCTGTGCTGATGATGATGCTGCTTAAAGTCCGCAAGAGCGCA 1101  
 QY 1078 ACCCTTGTCAATTAATTTGCAATCATTTGTGTTGGGCACTTTAATGAGATGCGGCTGA 1137  
 DB 1102 ACCCTTGTCAATTAATTTGCAATCATTTGTGTTGGGCACTTTAATGAGATGCGGCTGA 1161  
 QY 1138 CCGAGAGAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
 DB 1162 CCGAGAGAGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1221  
 QY 1198 TAAATCAATGCGCGGTACAGAGGCTGTCACACCGCGAGGGAGCTAATCTAGAAAGC 1257  
 DB 1222 TCAATCAATGCGCGGTACAGAGGCTGTCACACCGCGAGGGAGCTAATCTAGAAAGC 1281  
 QY 1258 GCGTGTAGTCCGATCGGATCTGCAACTGCACTCGGTGAAGTGGAAATGCTAGTAAT 1317  
 DB 1282 GCGTGTAGTCCGATCGGATCTGCAACTGCACTCGGTGAAGTGGAAATGCTAGTAAT 1341  
 QY 1318 CCGGATGACATGCTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377  
 DB 1342 CCGGATGACATGCTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1401  
 QY 1378 CATGGAGTGGGTTTACACAGAGCAGGTAGTCTAACCG 1416  
 DB 1402 CATGGAGTGGGTTTACACAGAGCAGGTAGTCTAACCG 1440

Search completed: May 5, 2006, 14:02:54  
 Job time : 938 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 5, 2006, 13:14:31 ; Search time 6222 Seconds  
(without alignments)  
10956.088 Million cell updates/sec

Title: US-10-659-980A-1  
Perfect score: 1457  
Sequence: 1 atgacgcctgscgcatcgc.....tcgagatcactgactggggtg 1457

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 70 summaries

Database :  
1: gb\_ebcl1:\*  
2: gb\_esc2:\*  
3: gb\_esc3:\*  
4: gb\_hnc1:\*  
5: gb\_esc4:\*  
6: gb\_esc5:\*  
7: gb\_esc6:\*  
8: gb\_esc7:\*  
9: gb\_gse1:\*  
10: gb\_gse2:\*  
11: gb\_gse3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	813.8	55.9	6499	9	BH771024	LMGtag74
2	720.4	49.4	1954	8	CX109203	RECM0593
3	648.4	44.5	1143	9	BH814966	PBPC15 Ma
4	641.4	44.0	1084	9	BZ439740	BONDUS4TF
5	637.8	43.8	941	7	CK296497	EST759211
6	633.2	43.5	1053	9	BZ447500	BONKFA0TR
7	627.8	43.1	1010	9	BZ431523	BONKHA87TF
8	622.2	42.7	1079	9	BH705272	BONKKA47TF
9	621.8	42.7	1016	9	BZ426201	BONDZ28TF
10	617.4	42.4	1031	9	BH656222	BH645222 BOMJ286TF
11	615.2	42.2	1077	9	BZ450751	BONBO63TR
12	613.2	42.1	1044	9	BZ463550	BONKCS6TF
13	613.2	42.1	1051	9	BZ499182	BONHR54TR
14	601.4	41.3	889	7	CK280527	EST743249
15	597.4	41.0	884	7	CK280528	EST743250
16	596.2	40.9	1054	9	BH647750	BONND15TR
17	592	40.6	974	9	BZ685785	PUBDT54TD
18	587	40.3	974	9	BZ440868	BONKTS1TR
19	582.2	40.0	756	8	DN469396	USDA-FP-1
20	574	39.4	787	7	CO159272	FLD1-12-C
21	574	39.4	951	9	BH651765	BONPES30TR
22	572.6	39.3	851	7	CK290947	EST753661

23	566.4	38.9	746	9	AQ957362	LERAP36TF
24	564	38.7	860	7	CN757363	CN757363 ID0AAACE
25	562.2	38.6	977	9	BZ459292	BONKQ28TF
26	560.4	38.5	947	9	BZ464837	BZ464837 BONKP76TR
27	559.4	38.4	901	9	BZ461852	BZ461852 BONKQ15TF
28	556.4	38.2	1106	9	BZ469058	BONNIE54TF
29	556.2	38.2	1069	9	BZ474941	BZ474941 BONKRS36TR
30	554.4	38.1	716	9	AQ957363	AQ957363 LERAP36TR
31	553.6	38.0	918	9	BZ437960	BZ437960 BONNV72TF
32	547.4	37.6	915	7	CV287759	aoF01-13m
33	547.4	37.5	704	3	BZ79543	BZ79543
34	542.2	37.2	1096	7	CV292617	aoF01-2m8
35	540.2	37.1	875	7	BZ502068	BONIT125TR
36	532.2	36.5	787	8	DN468747	DN468747 USDA-FP-1
37	526.2	36.1	862	9	BH562180	BH562180 BOGVN36TR
38	526.2	36.1	846	9	BH705306	BH705306 BOMIY08TF
39	524.8	36.0	873	9	BH705306	BH705306 BOGH290TF
40	524.6	36.0	875	9	BH547523	BZ426275 BOOAB18TR
41	524.4	36.0	889	9	BZ426275	CF566443 1115107G1
42	521.8	35.8	745	6	CF566443	BH482629 BOGNK19TR
43	521.6	35.8	842	9	BH482629	BZ483676 BOOAS16TR
44	521.4	35.8	837	9	BZ483676	CF557211 1115025D0
45	520.6	35.7	770	6	CF557211	BH565955 BOGIX59TF
46	519.2	35.6	866	9	BH565955	BH545929 BOGNL60TF
47	518.8	35.6	849	9	BH545929	BH700021 BOMND50TR
48	518.4	35.6	995	9	BH685417	BH685417 BOMGL56TF
49	517.6	35.5	1020	9	BH730827	BH730827 BOMR515TR
50	517	35.5	848	9	BH730827	BH501502 BONDSE57TR
51	516.6	35.5	848	9	BH535640	BH552149 BOHFE72TR
52	516.2	35.4	831	9	BZ501502	CM001852 ZMBH9000
53	516.2	35.4	835	9	BH552149	CM001852 ZMBH9000
54	515.6	35.4	1050	10	BH578853	BH578853 BOHOB24TR
55	515	35.3	857	9	BH704946	BH704946 BOMLQ71TF
56	514.8	35.3	845	9	BH473870	BH473870 BOHQA27TR
57	514.8	35.3	839	9	DR027023	DR027023 Osmo01058
58	514.8	35.3	895	8	BZ470443	BONR251TR
59	513.6	35.3	862	9	BZ470443	CV291378 aoF01-1m8
60	513	35.2	837	7	CV291378	BZ446240 BONJI04TR
61	512.8	35.2	822	9	BZ446240	BZ507023 BONIT76TR
62	512.8	35.2	824	9	BZ507023	BH529859 BOGOD80TR
63	512.8	35.2	837	9	BZ435283	BH529859 BOGOD80TR
64	512	35.1	862	9	BZ435283	BH529859 BOGOD80TR
65	512	35.1	1030	9	BZ463792	BZ463792 BONKY05TR
66	512	35.1	1082	9	BZ459110	BZ459110 BONBR59TR
67	511.4	35.1	819	9	BH484054	BH484054 BOHGA75TR
68	510.8	35.1	861	9	BH582045	BH582045 BONGU46TF
69	509.4	35.0	935	9	BZ468986	BZ468986 BONKD08TF
70	508.8	34.9	861	9	BH597210	BOHEB62TR

## ALIGNMENTS

RESULT 1  
BH771024 6499 bp DNA linear GSS 01-MAY-2002  
LMGtag746 MG1363 Random Sequence Tag Library Lactococcus lactis  
subsp. cremoris genomic, genomic survey sequence.

LOCUS  
BH771024  
DEFINITION  
LMGtag746 MG1363 Random Sequence Tag Library Lactococcus lactis  
subsp. cremoris genomic, genomic survey sequence.  
VERSION  
BH771024.1 GI:20373981  
KEYWORDS  
GSS.  
SOURCE  
Lactococcus lactis subsp. cremoris  
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
Lactococcus.

REFERENCE  
1 (bases 1 to 6499)  
Biolotin, A., Rhtlich, S. D. and Sorokin, A.  
Studies of genomes of dairy bacteria Lactococcus lactis  
Sci. Aliments (2002) In press  
Contact: Sorokin A  
Genetique Microbienne  
INRA  
CNU INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France

Tel: 33 1 34 65 25 16  
 Fax: 33 1 34 65 25 21  
 Email: borokine@jouy.inra.fr  
 Best homologue in strain IL1403 is ywga (78%)  
 Class: shotgun  
 High quality sequence start: 30  
 High quality sequence stop: 6471.  
 Location/Qualifiers

## FEATURES

Source

1.6499  
 /organism="Lactococcus lactis subsp. cremoris"  
 /mol\_type="genomic DNA"  
 /strain="MG1363"  
 /sub\_species="cremoris"  
 /db\_xref="taxon:1359"  
 /clone\_1ib="MG1363 Random Sequence Tag Library"  
 /note="Vector: pSGM02; Site\_1: SmaI; Library of  
 chromosomal fragments of L.lactis strain MG1363 was  
 prepared by partial AluI digestion or by sonication."

## ORIGIN

Query Match 55.9%; Score 813.8; DB 9; Length 6499;  
 Best Local Similarity 75.4%; Pred. No. 1e-239;  
 Matches 1106; Conservative 0; Mismatches 347; Indels 14; Gaps 7;

4 GAACGCTGGCGGCATGCTTTACATGCAAGTCGAACG--GCAGCAGGATGCTTGACATC 61  
 555 GAACCTGGCGGCGGCTGCTTAATGCAAGTTGAGCCATGAAAGATGCTTGACACCA 614  
 62 TGGTGGCAGTGGCGGAGCGGTGAGTATGAT-CGGAACGTATCCAGAAAGGGGGGGA 120  
 615 ATTTGAGAGCGAGCAACGGGTGAGTAAAGCGGTGGGAATCTGCTTTGAGCGGGGACA 674  
 121 AGCATCGAAAGATGCTTAATCCGCAT------TACTTAAGAGAAAGCAAGGG 173  
 675 ACATTGGAAAGAAATGCTTAATCCGCATTAATTAACATTAAGTTTAAAGTTTAA 734  
 174 ATCGAAGACCTTGGCTTTTGGAGCGGCGCATGCTGATTAAGTATGTTGGGGTAA 233  
 735 AGATCAATTTGATCATCTCAAGATGATCCCGCTTGTATTAAGTTGGTGAAGTAA 794  
 234 GGCCTACCAAGGCGAGCATCACTAGTGTCTGAGAGAGCAACGACCACTGGGACTG 293  
 795 GGCCTACCAAGGCGAGCATCACTAGTGTCTGAGAGAGTATCGGCGCATTTGGGACTG 854  
 294 AGACACGCGCCGAGATCTCTAAGGAGAGCGAGCGAGGAAATTTGACATATGGCGCA 353  
 855 AGACACGCGCCGAGATCTCTAAGGAGAGCGAGCGAGGAAATTTGACATATGGCGCA 914  
 354 CCTGATCCAGCAATCGCGGTGAGTGAAGAGCGCTTGGGTTGTAAGCTCTTCAATC 413  
 915 TCTGACCGAGCAACCGCGGTGAGTGAAGAGCGCTTGGGTTGTAAGCTCTTCAATC 974  
 414 GAGAAAGAAAG-GTTACGTTAAATATCTGATCTGATCGATCGATCGAAGAGAAC 472  
 975 GAGAAAGAAAGCTTGTGAGAGTGAAGAGCTCATCACTGACGCTTAATCCAGAAAG 1034  
 473 CGGCTTAATCTAGTCCAGAGACCGCGGTATATCTGAGGAGTGAAGCGCTTAATCG 532  
 1035 CGGCTTAATCTAGTCCAGAGACCGCGGTATATCTGAGGAGTGAAGCGCTTAATCG 1094  
 533 CTGGGCTTAAGAGGAGCGAGCGCTTGTAAAGTCAGATGTAAGTCCCGGGCTTAC 592  
 1095 TTGGGCTTAAGAGGAGCGAGCGCTTGTAAAGTCTGATGTAAGTCCCGGGCTTAC 1154  
 593 CTGGGCTTAAGAGGAGCGAGCGCTTGTAAAGTCAGATGTAAGTCCCGGGCTTAC 652  
 1155 CATTTGA-TGCATTGGAAGTGTGATCTGATGTCAGAGAGGAGAGTGAATTCATG 1213  
 653 TGTAGCAGTGAAGTGTGAGATATGGAAGAACTCATGTCGGAAGCGCTTCTGGG 712  
 1214 TGTAGCAGTGAAGTGTGAGATATGGAAGAACTCATGTCGGAAGCGCTTCTGGG 1273  
 713 TTAACACTGAGGCTCATGCAAGAAAGCGTGGGAGCAACAGGATTAATGATCCCTGG 772

1274 CTGTAACTGACACTGAGCTGAAAGCGTGGGAGCAAAACGATTAATACCTCGTAG 1333  
 773 TCACGCGCTTAACGATGTCACTAGTGTGGGCTTAATAG-GCTTGTATAGAACT 831  
 1334 TCACGCGCTTAACGATGTCACTAGTGTGGGAGCTTAATAGTCTTGTATGCACT 1393  
 832 AACGCGTGAAGTGAACGCTGAGGAGTACGCTGCAAGATTAATACTCAAGAACTGA 891  
 1394 AACGCGTGAAGTGAACGCTGAGGAGTACGCTGCAAGATTAATACTCAAGAACTGA 1453  
 892 CGGGAGCCCGCAACAGCGGTGATTAATGATTAATTCATGCAACGCGAAGAACTTGA 951  
 1454 CGGGAGCCCGCAACAGCGGTGATTAATGATTAATTCATGCAACGCGAAGAACTTGA 1513  
 952 CTAACCTTGAAT-ATGAGCAATTTTGAAGATATTTAGTGTGGGAGCGCTTAAC 1010  
 1514 CAGGCTTGAATATCTGCTATCTTCTTAAGATGAGAACTTCTTGGGAGCGGAT 1573  
 1011 AAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1070  
 1574 AAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1633  
 1071 GAGCGCAACCTTGTCAATTAATGCAATCATTTGGTGGGACTTAATGAGACTGCGG 1130  
 1634 GAGCGCAACCTTGTCAATTAATGCAATCATTTGGTGGGACTTAATGAGACTGCGG 1693  
 1131 TGACAAACCGAGAGAGGTGGGATGAGTCAATCTCATGCGCTTAATGAGAGCT 1190  
 1694 TGATTAACCGAGAGAGGTGGGATGAGTCAATCTCATGCGCTTAATGAGAGCT 1753  
 1191 TCACAGCTTAATCAATGCGGCTGCAAGAGGTGCAACCGCGAGGGAGCTAATCTC 1250  
 1754 ACACAGTGTCAATGAGATGTAACAGAGTCCGAGACAGTATGTTTAACTAATCTC 1813  
 1251 AGAAAGCGGCTGTAATGCTGAGTCCGATCGAATCTGCAATCTGCTGTAAGTCCG 1310  
 1814 TTAACCAATCTCTGATGCTGATGTAAGTGGGCAACTGCTTAATGAGTCCGATCTC 1873  
 1311 TAGTAATGCGGATGAGTATGCGGTGTAATACGTTCCGGGCTTGTATACACCGCC 1370  
 1874 TAGTAATGCGGATGAGTATGCGGTGTAATACGTTCCGGGCTTGTATACACCGCC 1933  
 1371 GTCAACCATGAGGAGGTGCTTCAACAGAGCAGTGTCTAACGTAAGAGGAGCGCTT 1430  
 1934 GTCAACCATGAGGAGGTGAGTATCCGAGATAGGTTGCTTAACGCAAGAGGAGCGCTT 1993  
 1431 GCCACGCTGAGATTCATGACTGGGCTG 1457  
 1994 CTTAAGTAAAGACGAGTACTGGGCTG 2020

RESULT 2  
 CX109203 1954 bp mRNA linear EST 03-JUN-2005  
 LOCUS CX109203  
 DEFINITION RCHM0593 A normalized whole-life-cycle cDNA library of rice *Oryza sativa* (indica cultivar-group) cDNA clone E1051022, E105308, E1008B17, B1076119, B1154A06, B113B10, E111 5', mRNA sequence.  
 ACCESSION CX109203  
 VERSION CX109203.1 GI:66922355  
 KEYWORDS EST.  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM *Oryza sativa* (indica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzoideae; *Oryza*.  
 1 (bases 1 to 1954)  
 REFERENCES  
 Zhang, J., Feng, Q., Jin, C., Qiu, D., Zhang, L., Xie, K., Yuan, D., Han, B., Zhang, Q. and Wang, S.  
 Features of the expressed sequences revealed by a large-scale analysis of ESTs from a normalized cDNA library of the elite indica rice cultivar Minghui 63  
 Plant J. 42 (5), 772-780 (2005)

JOURNAL

PUBMED  
Contact: Wang S  
National Key Laboratory of Crop Genetic Improvement  
Hauzhong Agricultural University  
Wuhan 430070, China  
Tel: 86-27-87282044  
Fax: 86-27-87287092  
Email: shijingwang@hotmail.com  
Seq primer: 77.

FEATURES  
source  
location/Qualifiers  
1..1954  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="mRNA"  
/strain="indica"  
/cultivar="Minghui 63"  
/db\_xref="taxon:39946"  
/clone="E1051022, E1053J08, E1008B17, B1076L19, B1154A06, B1133B10, E111"  
/tissue\_type="whole plant"  
/dev\_stage="whole-life-cycle"  
/lab\_host="E. coli DH10B"  
/clone\_lib="A normalized whole-life-cycle cDNA library of rice"  
/note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; The library is constructed based on the strategy of saturation hybridization with genomic DNA using rice cultivar Minghui 63. This library consists of cDNA from 15 directionally cloned cDNA libraries constructed with different tissues from 9 developmental stages."

ORIGIN  
Query Match 49.4%; Score 720.4; DB 8; Length 1954;  
Best Local Similarity 73.1%; Pred. No. 5.9e-211;  
Matches 1069; Conservative 0; Mismatches 336; Indels 57; Gaps 9;

3 TGAACGCTGCGCGCATGCTTTACATGCAATGCAAGCGGACGAGATCTTGCATCT 62  
130 TGAACGCTGCGCGCATGCTTTACATGCAATGCAAGCGGAGTGGT----- 177  
63 GGTGGGAGTGGCGGACCGGTGAGTAATGATCGGAACGTATCCAGAAAGGGGGTAA 122  
178 -GTTTCAAGTGGGAAACGGGTGAGTAACGCTGTAAGAACTGCTTGGGGGAAACAC 236  
123 GCATCGAAGATGCTGTAATACCGCATATCTCTAAGGAGAAAGCGGAGTCAAGAA 182  
237 AATCGAAGATGCTGTAATACCGCATATCTCTAAGGAGAAAGCGGAGTCAAGAA 283  
183 CCTTGGGCTTTTGAAGCGCGCATGCTGTAATGCTGTAAGTGGGGTAAAGCTTACCA 242  
284 AATCGGCGCAAGAGAGGGCTCGGTGATGATGCTGTAAGTGGGGTAAAGCTTACCA 343  
243 AGCGGACATCAATGCTTGGTCTGAGAGACACACGACCACTGGGATCGAGACCGG 302  
344 AGCGGATGATCAATGCTTGGTCTGAGAGACACACGACCACTGGGATCGAGACCGG 403  
303 CCAAGCTCTTCAAGGAGGACGAGTGGGAAATTTGACATGGGCGCAAGCTGTATCA 362  
404 CCAAGCTCTTCAAGGAGGACGAGTGGGAAATTTTCCGCAATGGGCGCAAGCTGTATCA 463  
363 GCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGTAAGCTTTTCAAGTGAAGAA 422  
464 GCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGTAAGCTTTTCCGAGAAAGAA 523  
423 AGTTACGTTAATATCGTCACTGATGACGCTTTCAGACAGAAAGCAACCGGTACTA 482  
524 -----CAATGACGCTTTCAGACAGAAATTAACATCGGTACTA 560  
483 CGTGCAGACGCGCGGTAAATCAATGAGGTCAGAGCTTATCGAATTAATCGGCGCTAA 542  
561 TGTCCAGACGCGCGGTAAATCAATGAGGTCAGAGCTTATCGAATTAATCGGCGCTAA 620  
543 AGGTGCGCAGCGCGCTTTTGAATGATGAAATCCCGGCTTAACTGGAATTG 602

621 AGCGTGTAGTGTGCTTTTCAAGTCCGCGCTCAAAATCCAGGAGCTCAACCTTGACAG 680  
603 GCTTTGAAATCAACAGCTAGTGTGACAGAGGAGTGGAAATTCATGTGTAGTGTG 662  
681 CGTGGAAATCAACAGCTAGTGTGACAGAGGAGTGGAAATTCATGTGTAGTGTG 740  
663 AATGCGTGTAGTGTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 722  
741 AATGCGTGTAGTGTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 800  
723 CGTGTATGACAGAAAGCTGGGAGAGCAACAGATTTAGTATCCTCG--TGTCCAGCC 780  
801 CACTGAGAGACGAAAGCTTGGGAGAGCAATATGATGAGAGCCCAAGTATTCATACC 860  
781 CTAAACGATGTCACTAGTGTGGGCTTATTTAGGCTTG--GTACGAGCTAACCGGT 838  
861 GTAAACGATGTCACTAGTGTGGGCTTATTTAGGCTTG--GTACGAGCTAACCGGT 920  
839 GAAGTTGACCGGCTGGGAGTACCGGTGCAAGATTTAAATCAAGAAATGACCGGGAC 898  
921 TAAATATCCCGCTGGGAGTACCGGTGCAAGATTTAAATCAAGAAATGACCGGGAC 980  
899 CCGCAACAGCGGTGATTTATGATTTATTTGATGCAACCGGAAATCTTACTACC 958  
981 CCGCAACAGCGGTGATTTATGATTTATTTGATGCAACCGGAAATCTTACTACC 1040  
959 TTGACATGTAGCAATTTTCTTGAATGATTTATTTG--CTTGGGAAACGCTAACAGAG 1017  
1041 TTGACATGTAGCAATTTTCTTGAATGATTTATTTG--CTTGGGAAACGCTAACAGAG 1100  
1018 CTGCAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077  
1101 GTGCAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1160  
1078 ACCCTTGTATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1137  
1161 ACCCTTGTATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1219  
1138 CCGGAGAAAGTGGGAGTACGTAATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 1197  
1220 CCGGAGAAAGTGGGAGTACGTAATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 1279  
1198 TAAATCAATGCGGCTGTAACAGAGGCTTCCAAACCGGAGGAGGAGCTTATCTCA 1257  
1280 TGTCAATGCGGCTGTAACAGAGGCTTCCAAACCGGAGGAGGAGCTTATCTCA 1339  
1258 GCGTGTAGTCCGATCGGAGTCTGCACTCGACTCGTGAAGTGGATGCTTATTAAT 1317  
1340 GGTGTGTAGTCCGATCGGAGTCTGCACTCGACTCGTGAAGTGGATGCTTATTAAT 1399  
1318 CCGGATCAG-CATGTCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCCTCA 1376  
1400 CGCGGTCAGCATACGAGGCTGAATCCGTTCCCGGCTTTGTACACACCGCCCTCA 1459  
1377 CCAATGAGATGAGTTTCAACAGAGGATGTC-TAACGTAAGAGGAGGCTTCCAC 1435  
1460 CTATAGAACTGGCCATGTTTGAATCATTAACCTTAACCTTAAGAGGAGGATGCTTAA 1519  
1436 GGTGAGATTCATGACTGGGCTG 1457  
1520 GGTGAGATTCATGACTGGGCTG 1541

RESULT 3  
BH814966/c 1143 bp DNA linear GSS 31-MAY-2003  
LOCUS BH814966  
DEFINITION BPHC15, Maize B chromosome PCR DNA library Zea mays genomic c clone  
ACCESSION BH814966  
VERSION BH814966.1 GI:31249923  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
AUTHORS  
TITLE  
Cheng Y.M. and Lin B.Y.

JOURNAL  
PUBMED  
Cloning and characterization of maize B chromosome sequences derived from microdissection  
Genetics 164 (1), 299-310 (2003)

COMMENT  
12750341

Contact: Bor-yaw Lin  
Department of Molecular Biology  
National Chung Hsing University  
250 Kuo Kuang Rd, Taichung 402, Taiwan (ROC)  
Tel: 886-4-2285-1885  
Fax: 886-4-2287-4879  
Email: bylin@dragon.nchu.edu.tw

Insert Length: 1143 Std Error: 0.00  
Classes: PCR fragment.

FEATURES  
Location/Qualifiers

1..1143  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="L289"  
/db\_xref="taxon:4577"  
/clone="PBPC15"  
/sex="hermaphrodite"  
/cell\_type="Pollen mother cells"  
/dev\_stage="Meiosis I"  
/lab\_host="DH5alpha"  
/clone\_1ib="Maize B chromosome PCR DNA library"  
/note="Vector: pBSK-; The library was constructed from microdissected B chromosome DNA."

ORIGIN

Query Match 44.5%; Score 648.4; DB 9; Length 1143;  
Best Local Similarity 89.6%; Pred. No. 9,9e-189;  
Matches 744; Conservative 0; Mismatches 76; Indels 10; Gaps 4;  
631 CAGAGGAGGTGAATTCATGTGTGACGTGAATGCTTACAGATATGGAAGACATCG 690  
1143 CAGAGGGGGGTGAATTCACCTGTGACGTGAATGCTTATATGTGGAGAACACCG 1084  
691 ATGGGGAAGGAGCGCTCTGGGTGAACACTGACGCTCAAGCAAGCGTGGAGCA 750  
1083 ATGGGGAAGGAGCGCTCTGGGTGAACAAGTTCATGCTCAAGCAAGCGTGGAGCA 1024  
751 ACAGATTAGATACCTGTGTGACGCGCTTAAAGATGTCACTAGTTGTGGCGCTT 810  
1023 ACAGATTAGATACCTGTGTGACGCGCTTAAAGATGTCTACTAGTTGTGGCGCTT 964  
811 ATTAGGCTTGTGATCAAGCTTAAGCGGTGAAGTTGACGCGCTGGGAGTACCGTGCAG 870  
963 AATTGACGTTGTGAACGAGCTTAACCGGTGAATTAACCGCTGGGAGTACCGTGCAG 904  
871 ATTAAACCTCAAGGAATTGACGGGAGCCCGCAAGGGGTGATTATGTGATTAATTC 930  
903 ATTAAACCTCAAGGAATTGACGGGAGCCCGCAAGGGGTGATTATGTGATTAATTC 844  
931 GATGCAAGCGCAAAACCTTACCTTACCTTGCATGTAGCGAATTTCTAGAGATAGATT 990  
843 GATGCAAGCGCAAAACCTTACCTTACCTTGCATGTAGCGAATTTCTAGAGATAGATT 784  
991 AGTGCT---TGGGGAACGCTTAACAGGTGCTGATGGCTGTGTCAGTCTGTCTGCA 1047  
783 AGTGCTCAAAAGAGAACAGTACACAGGTGCTGATGGCTGTGTCAGTCTGTCTGCA 724  
1048 GATGTGGGTTAAGTCCCGCAACGAGCGCAACCTTGTCACTTAATTGCACTATTGGTT 1107  
723 GATGTT--GTTTATGCTCCGCAACGAGCGCAACCTTGTCACTTAATTGTTGTCAGAAA----- 670  
1108 GGGCACTTAAATGAGACTGCGGTGACAAACCGAGAGAGGTGGAGATGACGTCAAGTCC 1167  
669 GGGCACTTAAATGAGACTGCGGTGACAAACCGAGAGAGGTGGAGAT--ACGTCAAGTCC 611

QY 1168 TCATGGCCCTTATGGGGTGGGCTTACACGTAAATACAAATGGCGGCTACAGAGGGTTGCCA 1227  
DB 610 TCATGGCCCTTATGGGGTGGGCTTACACGTAAATACAAATGGCGGCTACAGAGGGTTGCCA 551  
QY 1228 ACCCGGAGGGGAGACTAATCTCAGAAACCGCGTGTAGTCCGAGTCGGAGTCCGCACT 1287  
DB 550 ACCCGGAGGGGAGACTAATCTCAGAAACCGCGTGTAGTCCGAGTCCGAGTCCGCACT 491  
QY 1288 GCACTCCGTGAAGTTCGGAATGCTTATGTAATGCGGATATGACATGTCCGGGTGAATACGTT 1347  
DB 490 CGACTGCATGAAGTTCGGAATGCTTATGTAATGCGGATATGACATGTCCGGGTGAATACGTT 431  
QY 1348 CCCGGGCTTGTATACACACGCCCGGTACACCATGGAGTGGGTTTACACAGAGCGGTA 1407  
DB 430 CCCGGGCTTGTATACACACGCCCGGTACACCATGGAGTGGGTTTACACAGAGCGGTA 371  
QY 1408 GTCTTAACCTTAAGAGGCGGCTTGCACAGTGAATTCATGACTGGGGTGG 1457  
DB 370 GTCTTAACCTTAAGAGGCGGCTTGCACAGTGAATTCATGACTGGGGTGG 321

RESULT 4  
BZ439740 1084 bp DNA linear GSS 13-DEC-2002  
LOCUS BZ439740  
DEFINITION BONDUS4TF BO.1.6.2\_KB tot Brassica oleracea genomic clone BONDUS4,  
genomic survey sequence.  
ACCESSION BZ439740  
VERSION BZ439740.1 GI:26695676  
KEYWORDS GSS.  
SOURCE  
ORGANISM  
Brassica oleracea  
Brassica oleracea

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 1084)  
Ayle, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Utehrack, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)

COMMENT  
Other GSSs: BONDUS4TR  
Contact: Chris Town

TIER 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TP  
Class: sheared ends.

FEATURES  
Location/Qualifiers

1..1084  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="T01000DH3"  
/db\_xref="taxon:3712"  
/clone="BONDUS4"  
/clone\_1ib="BO.1.6.2\_KB tot"  
/note="Vector: pHOS1; Site 1: BstXI; 1.6-2 kb sheared total DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Query Match 44.0%; Score 641.4; DB 9; Length 1084;  
Best Local Similarity 77.2%; Pred. No. 1.4e-186;  
Matches 841; Conservative 0; Mismatches 221; Indels 27; Gaps 4;  
237 CTACCAAGGCGACATCATGATGTTGTTCTGAGAGAGACCAAGCACTGGAGCTGAGA 296  
DB 3 CTACCAAGGCGACATCATGATGTTGTTCTGAGAGAGACCAAGCACTGGAGCTGAGA 61  
QY 297 CAGGCGCAAGACTCTTACCGGAGGACAGAGTGGGAATTTTGGCAATYGGGCGCAAGCT 356

```

Db      |||||
62  CACGCGCCAGACTCTCTACCGGAGGAGCAGTGGGGAATTTTCCGCAATGGGGGAAGCTT 121
Qy      |||||
357 GATTCAGCAATGCGCGTGAAGGAGGAGGCTTGGGGTTGTAAGCTCTTTCAGTCAAG 416
Db      |||||
122 GACGAGCAATGCCCGGTGAGAGTGAAGAGGCTTACGGGTCTGAACTTTTCCCAAG 181
Qy      |||||
417 AAGAAAGGTTACGGTAAATATCGTACTGACGCTATGACAGAGAGCAACCGCGC 476
Db      |||||
182 AAGA-----ACATATGACGATATCGGGGAATAGATCGGC 218
Qy      |||||
477 TAACACTGTCGACAGACCGCGTAATACGTAAGGTCGCAAGCTTATGGAATTAATCG 536
Db      |||||
219 TAACCTGTGTCGACAGACCGCGTAATACGTAAGGTCGCAAGCTTATGGAATTAATCG 278
Qy      |||||
537 GCGTAAAGGTCGACAGCGGCTTTGTAGTCAAGTGAATCCCGGGCTTAACCTCG 596
Db      |||||
279 GCGTAAAGGTCGTAAGTGGCTTTTAAGTCCGCGCTCAAAATCCAGGGCTCAACCTCG 338
Qy      |||||
597 GAATTCGCTTGAATCTACAGGCTAGAGTGGCAGAGGAGGTGAAATTCATGTGTA 656
Db      |||||
339 GACAGCGGTGAAACTACAGCTTGAATGAGTGGGAGAGGGAATTTCCGGTGA 398
Qy      |||||
657 GCAATGAAATGCGTAGAGATGAAAGAAATCATGATGCGAAAGCAGCTCTCGGTTAA 716
Db      |||||
399 GCGGTAAATGCGTAGAGATGAAAGAAACCAACGCGGAAAGCACTCTGCTGGCCGA 458
Qy      |||||
717 CACTGACGCTACGACGGAAGCGTGGGAGCAAAAGATTAAGATACCTGGTAGTCA 776
Db      |||||
459 CACTGACGCTACGAGAGCAAAAGCTAGGGAGCAAAAGGATTAAGATACCTGGTAGTCA 518
Qy      |||||
777 CGCCCTAAACGATGTCATAGTGTGGGCTTATTAAGGCTTG--GTAACGAAGCTAAC 834
Db      |||||
519 AGCCGTAACGATGATTAAGTGGGCTGTGCGTATGACCCCGTGCAGTGTGATGCTAAC 578
Qy      |||||
835 GCGTGAAGTGAACCGCTGCGGAGATACGCTGCGAAGATTAACCTCAAGATGAAG 894
Db      |||||
579 GCGTGAAGTGAACCGCTGCGGAGATACGCTGCGAAGATTAACCTCAAGATGAAG 638
Qy      |||||
895 GGAACCGGCAAGCGGTGATTAATGAGTAAATGAGTGAACGGAAGAAACCTTACCT 954
Db      |||||
639 GGGCCCGCAAGCGGTGAGAGTGTGTTAATGAGTGAAGGAGAAAGAACCTTACCA 698
Qy      |||||
955 ACCCTGACATGATGAGGAATTTTCTAGAGATGATTAATG-CTTCCGGAACGCTTAACA 1013
Db      |||||
699 GGGCTTGAATGCGCGGAATCTCTTGAAGAGAGGGGTGCTTCCGGGAACGCGGACAC 758
Qy      |||||
1014 GGTGCTGCAATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1073
Db      |||||
759 GGTGCTGCAATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
Qy      |||||
1074 CGCAACCTTGTGATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
Db      |||||
819 CGCAACCTTGTGATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 878
Qy      |||||
1134 CAACCGGAGAGAGTGGGAGTGAAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1193
Db      |||||
879 TAAGCGGAGAGAGTGGGAGTGAAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 938
Qy      |||||
1194 CACGTAATACATGAGCGCTGACAGAGGTTGCCAACCGCGGAGGAGGAGTAAATCTTACA 1253
Db      |||||
939 CACGTAATACATGAGCGCTGACAGAGGTTGCCAACCGCGGAGGAGGAGTAAATCTTACA 998
Qy      |||||
1254 AAGCGGCTGCTAGTCCGAGTCCGAGTCTGCACTGCACTCCGTAAGTCCGAATCCGTAAG 1313
Db      |||||
999 AAGCGGCTGCTAGTCCGAGTCCGAGTCTGCACTGCACTCCGTAAGTCCGAATCCGTAAG 1058
Qy      |||||
1314 TAATCGCG 1322
Db      |||||
1059 TAATCGCG 1067

```

```

CK296497 941 bp mRNA linear EST 02-AUG-2004
LOCUS CK296497
DEFINITION EST75211 Nicotiana benthamiana mixed tissue cDNA library, NBM226 5'
normalised, full-length Nicotiana benthamiana cDNA clone NBM226 5'
end, mRNA sequence.
ACCESSION CK296497
VERSION CK296497.1 GI:39881943
KEYWORDS EST.
SOURCE Nicotiana benthamiana
ORGANISM Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 941)
Staskiewicz, B., Jin, H. and Baker, B.
Buell, C.R., Hart, A., Zismann, V., Karameycheva, S.A., Day, B.,
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
CONTACT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA TAG.
Location/Qualifiers
1..941
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBM226"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr; and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

```

## ORIGIN

Query Match 43.8%; Score 637.8; DB 7; Length 941;  
 Best Local Similarity 81.3%; Pred. No. 1.8e-185;  
 Matches 752; Conservative 0; Mismatches 172; Indels 1; Gaps 1;

```

Qy      1 ATTGAACGTCGGCGGATCTTACACATGACAGTGAACGGCAGACGAGATGCTGCAT 60
Db      |||||
Qy      18 ATTGAACGTCGGCGGATCTTACACATGACAGTGAACGGCAGACGAGATGCTGCAT 77
Db      |||||
Qy      61 CTGGTGGCAGTGGCGGAGCGGTGATGATCGGAACGTAATCCAGAGAGGGGGTGA 120
Db      |||||
78 CTGGTGGCAGTGGCGGAGCGGTGATGATGCTGAGATCTGCTGATGATGGGGGATA 137
Qy      |||||
121 ACGCATGAAGATGTGTAATACCGCATATATCTTAAGAGAGAAAGCAGGGAGTGA 180
Db      |||||
138 ACGCTGGAAGACGATTAATACCGCATATATCTTAAGAGAGAAAGCAGGGAGTGA 197
Qy      |||||
181 GACCTTGGCGCTTTGAGAGCGCGCATGCTGATTAAGTGTGTTGGTGGGTTAAAGGCTAC 240
Db      |||||
198 GGCCTTGGCGCTTCAATGAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 257
Qy      |||||
241 CAAGCGGAGATCAATGATGTTGTTGAGAGAGAGCAGACCACTGAGGACTGAGACAG 300
Db      |||||
258 CAAGCGGAGATCAATGATGTTGTTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 317

```

QY	301	GGCCGAGACTCCCTTAACGGGAGGAGCAGAGTGGGGAAATTTGGACAAATGGGGCGAAGCCTGATC	360
Db	318	GTCCGAGACTCTTAACGGGAGGAGCAGAGTGGGGAAATTTGGACAAATGGGGCGAAGCCTGATC	377
QY	361	CAGCAATGCGCGGTGATGATGAAGAGGCGCTTCGGGTGTGAAGCTCTTTCAGTCGAGAGA	420
Db	378	CAGCCATGCGCGGTGATGATGAAGAGGCGCTTCGGGTGTGAAGCTCTTTCAGTCGAGAGA	437
QY	421	AAAGGTTACGGTAAATATCTGTACTCATGACGGTATGCAAGAAAGAACCGGCTAAC	480
Db	438	AGGGCAGTTACTCTAATATACGTGATTTGTTTGGAGTTAACGACGAAATAGCACCGGCTAAC	497
QY	481	TACGCGCCAGCAGCGCGGCTAATAGTGGGGGTGCAAGCTTAATGGGAATTACTGGGGGT	540
Db	498	TCTGTGCCAGCAGCGCGGCTAATAGCAGGGGTGCAAGCTTAATGGGAATTACTGGGGGT	557
QY	541	AAAGGGTGGCGAGGGCGGCTTTGTAAAGTCAGATGTGAAATCCCGGGCTTAACCTGGGAAT	600
Db	558	AAAGGGCGCGGTGATGATGATTTGTTAAGTTGAATGTGAAATCCCGGGCTTAACCTGGGAAT	617
QY	601	TGCGTTTGAACCTACAGGCTTAGAGTGTGGCAGAGGGAGTGGAAATTCATGTGTAGCAG	660
Db	618	TGCATCCAAACTGGCAGAGCTAGATGTGTAGAGGGGTGTGGAATTTCTCTGTAGACGG	677
QY	661	TGAATATGCTTAAGATATATGGAGAATCATGATGCGGAGGCGAGCCTCTCGGGTTAACT	720
Db	678	TGAATATGCTTAATATATGGAGAATCATGATGCGGAGGCGAGCCTCTCGGGTTAACT	737
QY	721	GACGCTCATGACAGAAAGCGTGGGGAGCAAAACAGATTAGATACCTGTGATGCCAGCC	780
Db	738	GACACTGATGATGCGAAAGCGTGGGGAGCAAAACAGATTAGATACCTGTGATGCCAGCC	797
QY	781	CTAAACGATGTCACTAGTGTGTGGGCGCTTAATTAAGGCTTGTGAACGAAGCTAACCGGTGA	840
Db	798	GTAACGATGTCACTAGTGTGTGGGCGCTTAATTAAGGCTTGTGTAGTGGAGCTAACCATTA	856
QY	841	AGTTACACCGGCTTGGGGAGTACCGGTGGCAAGATTAAATCTCAAGAAATTGACGGGAGCC	900
Db	857	AGTTACACCGGCTTGGGGAGTACCGGCGCAGGTTAAATCTCAAGAAATTGACGGGAGCC	916
QY	901	GCACAAGCGGTGATTAATGTGATT	925
Db	917	GCACAAGCGGTGATTAATGTGATT	941
RESULT 6			
BZ447500/C			
LOCUS	BZ447500	1053 bp	DNA linear GSS 13-DEC-2002
DEFINITION	BONKRF4OTR BO.1.6.2_KB_tot Brassica oleracea genomic clone BONKRF40,		
ACCESSION	BZ447500		
VERSION	BZ447500.1		
KEYWORDS	GSS.		
SOURCE	Brassica oleracea		
ORGANISM	Brassica oleracea		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.		
AUTHORS	1 (bases 1 to 1053)		
TITLE	Ayale, M., Haas, B., J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Uteirbeck, T.R., Wortman, J.R., White, O.R. and Town, C.D.		
JOURNAL	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis		
PUBMED	Genome Res. 15 (4), 487-495 (2005)		
COMMENT	15805490		
	Other_GSS: BONKRF40TF		
	Contact: Chris Town		
	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA.		
	Tel: 301-838-3523		
	Fax: 301-838-0208		
	Email: cdtown@tigr.org		
	DNA is from a doubled haploid provided by Tom Osborn.		

FEATURES		Seg primer: TR
source	Location/Qualifiers	Class: sheared ends
1..1053	/organism="Brassica oleracea" /mol_type="genomic DNA" /strain="TO1000DH3" /db_xref="taxon:3712" /clone="BONKp40" /clone_1id="BO_1.6_2_KB_tot" /note="vector: pHOS1; Site 1: BstXI, 1.6-2 kb sheared total DNA inserted into pHOS1 using BstXI linkers"	
Query Match	43.5%; Score 613.2; DB 9; Length 1053;	
Best Local Similarity	77.2%; Pred. No. 4.9e-184;	
Matches 831, Conservative 0;	Mismatches 218; Indels 27; Gaps 4	
Db	200 GGGCGATGCTGATTAGCTAGTGTGGTGGGAAAGGCGCTTACCAAGCGGACGATCAGTAGT 259	
Db	1052 GAGTCGCGCTGATTAGCTAGTGGGTGGGCAATGCTTACCAAGCGGACGATCAGTAGTGC 993	
Qy	260 TGGTCTGAGAGACGACCGACCACTGGGACTGAGACACGCGCCAGACTCTCTACGGGAG 319	
Db	992 TGGTCGAGAGAGATGATCAGCCACACTGGGACTGAGACACGCGCCAGACTCTCTACGGGAG 933	
Qy	320 GAGCAGATGGGGAAATTTTGACAAATGGGCGCAAGCTGTATCCAGCAATATGCCCGCTGAGTG 379	
Db	932 GCAGCAGATGGGGAAATTTCCG-CAATGGGCGAAAGCTGTACGAGCAATATGCCCGCTGAGTG 874	
Qy	380 AAGAAAGGCTTCGGGTGTGAAGCTCTTCAAGTGGAGAAAGAAAGTTACGGTAAATATAT 439	
Db	873 TAGAAGGCTTACGGGTCTGAACTTCTTTCCAGAGAAAG----- 833	
Qy	440 CGTGACTATGACGGTATCGACAGAAAGACCGGCTTACATGATCGACAGCCGCGG 499	
Db	832 ----AGCAATACCGTATCTGGGGAAATGAGCATGGGCTTAACTGTGTCACAGCCGCGG 777	
Qy	500 TAAATACGATGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGGGTGGCAGCGCGCT 559	
Db	776 TAAATACAGAGATGCAAGCGTTATCCGGAATGATTTGGGCGTAAAGCGCTGTCTAGGTGACT 717	
Qy	550 TTGTAAAGCAATGTGAATATCCCGGGGTTAATCGGGAATTTGGGTTTGAACCTACAAAG 619	
Db	716 TTTTAAAGTCCCGCGTCAATCCACAGGGCTCAACCTGGACAGCGGTGGAATCTACCAAG 657	
Qy	620 CTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTAGCACTGAATTCGTAGAGATATG 679	
Db	656 CTTAGATACGTAAGGGGACAGAGGAATTTCCGGTGGAGCGGTGAATATCGTAGAGATGCG 597	
Qy	680 GAAAGAACATGATGGCGAAGGCGAGCTCTTGGGTTTAACTGACGCGCTCATSCAGAAAGC 739	
Db	596 AAAAGAACCAACAGCGCGAAACCACTGTCTGGGCGGACACTGACACTGAGAGACGAAAGC 537	
Qy	740 GTGGGAGCAACACAGATTAAATACCTCGTAAATGCCACGCGCTTAAAGATGCAATGAT 799	
Db	536 TAGGGAGACGAATGGGATTAAATACCCCAATGATCTAGCCGTAAACGATGATATCTAGG 477	
Qy	800 TGGTGGGCGCTTAATAGGCTTG--GTAAAGAGCTTAAACGCGTGAAGTTGACCGGCTGGGGA 857	
Db	476 CGCTGTGCTATCGACCCGTCGATGCTGTAGCTTAAACCGCTTAAAGTATCCCGCTGGGGA 417	
Qy	858 GTACGGTGCAGATTAAACTCAAGGAATTTGACGGGAGCCGACAAAGCGGTGATTTA 917	
Db	416 GTAGCTTGGCAAGATGAATCTCAAGGAATTTGACGGGAGCCGACAAAGCGGTGAGCA 357	
Qy	918 TGTGATTAATTCGATGCAACGCGAAATCTTACCTTGTGACATGTAGCGAATTTT 977	
Db	356 TGTGATTAATTCGATGCAAGGCAAGAACTTACCGGGCTTGAACATGACCGCAATCTT 297	
Qy	978 CTAGAGATAGATTAGTG-CTTGGGAAAGCTTAAACACAGGTCGTGATGGCTGTGCTAGC 1036	
Db	296 CTTGAAGAGAGGGGTGCTTGGGAACGCGGACACAGGTGTGATGCTGTGCTAGC 237	

QY 1037 TCGTGTCTGAGATGTTGGGTTAAAGTCCCGCAGACGAGCCAAACCTTGTCTATTAATGGC 1096  
DB 236 TCGTGTCTGAGATGTTGGGTTAAAGTCCCGCAGACGAGCCAAACCTTGTCTATTAATGGC 1177  
QY 1097 ATCATTTGGTTGGGCACTTTAATGAGACTGCGGGTGAACAACCGAGAGAGGTGGGGATG 1156  
DB 176 ACCGTTGAGTTTGAACCTGTAAACAGACTGCGGGTATAGCCGGAGAGAGGTGGGGATG 1117  
QY 1157 ACGTCAAGCTCTCATGAGCCCTTATGAGGTTAGGCTTACACATATATATATGCGGTACA 1216  
DB 116 ACGTCAAGCTCTCATGAGCCCTTATGAGGTTAGGCTTACACATATATATATGCGGTACA 57  
QY 1217 GAGGTTGCGCAACCGCGAGGCGAGCTAATCTCAGAAAACCGCTGTATGTCGGA 1272  
DB 56 AAGGTCGCGATCCAGAGAGGGTGAAGTAACTCCAAAACCGCTCTCAGTTCCGA 1

RESULT 7  
BZ431523 1010 bp DNA linear GSS 13-DEC-2002  
LOCUS BOMND88TF BO.1.6.2 KB lot Brassica oleracea genomic clone BOMND88,  
DEFINITION genomic survey sequence.  
ACCESSION BZ431523  
VERSION BZ431523.1 GI:26677909  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 1010)  
Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Utterback, T.R., Mortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490  
JOURNAL PUBMED  
COMMENT Other GSSs: BOMND88TR  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seg primer: TF  
Class: sheared ends.  
Location/Qualifiers  
1..1010  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOMND88"  
/note="Vector: pBIO51; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pBIO51 using BstXI linkers"

ORIGIN  
Query Match 43.1%; Score 627.8; DB 9; Length 1010;  
Best Local Similarity 77.9%; Pred. No. 2,3e-182;  
Matches 805; Conservative 0; Mismatches 202; Indels 26; Gaps 3;

QY 211 GATTAGCTAGTGGTGGGTTAAAGGCTTACCAAGGAGCATCATGTTGTTCTGAGAG 270  
DB 1010 GATTAGCTAGTGGTGGGTTAAAGGCTTACCAAGGAGCATCATGTTGTTCTGAGAG 951  
QY 271 GAGCAACAGCACAAGCTGGAGCTGAGACAGCGCCAGACTCTCTACGGAGGAGAGAGTGG 330  
DB 950 GATGATAGCACAAGCTGGAGCTGAGACAGCGCCAGACTCTCTACGGAGGAGAGAGTGG 891  
QY 331 GAATTTTGAACAATGGCGCAAGCGCTGATCCAGCAATGCGCGTGAAGAGAGGCTT 390

DB 890 GAATTTTCCGCAATGGGAGAAAGCTGACGAGCAATCCCGGTGAGGTAGAAAGCCCTA 831  
QY 391 CGGTTGTAAGCTCTTCACTGACGAGAAAGAAAGTTACGTTAATATCGTCACTCAG 450  
DB 830 CGGTTCTGAACCTCTTCTTCCAGAGAAAGAAAGTTAATATCGTCACTCAG 794  
QY 451 ACGGTATGACAGAAAGAAAGCAACCGGCTAACTACGTCAGAGCGCGGTAAATAGTAG 510  
DB 793 ACGGTATGAGAAATAGCAATCGCTAATCTGTGCGACAGACCGCGGTAAATAGTAG 734  
QY 511 GTGCAAGCTTAATCGAATTAATGAGGCTTAAAGGTCGAGCGCGCTTGTAAATGAC 570  
DB 733 ATGCAAGGTTAATCGAATGATGAGGCTTAAAGGCTCTGATAGGTCCTTTTAAAGTCG 674  
QY 571 ATGCAAAATCCCGGCTTAACTGAGAAATGCGTTTAAATCAAAAGCTTAAGTGG 630  
DB 673 CCGTCAAAATCCAGAGGCTCAACCTTGACAGCGGTGAAACTACCAAGCTTGAATACG 614  
QY 631 CAGAGGAGAGTGAATTCATGTTGAGAGTGAATGCGTAGAGATATGAAAGAACATCG 690  
DB 613 TAGGGGAGAGGAAATTTCCGTTGAGCGGTGAATGCGTAGAGATCGAAAGAACACCA 554  
QY 691 ATGGCAAGGAGCGCTCTGAGTTAACTGACGCTCACTGACAGAAAGCGTGGAGCA 750  
DB 553 ACGCGAAAGCACTGTGCTGGCCGACACTGACACTGAGAGACGAAAGCTAGGGAGCA 494  
QY 751 ACAGATTAAGATACCTGAGTGTGACAGCGCTTAAACGATGTCACTAGTTGTGGCCCT 810  
DB 493 ATGGATTAAGATACCTGAGTGTGACAGCGCTTAAACGATGTCACTAGTTGTGGCCCT 434  
QY 811 ATTAGGCTTG--GTAAAGAGCTAACGCGTGAAGTTGACCGCGTGGGAGTACGGTCGA 868  
DB 433 TCGACCGGTGACGTCGTGAGCTAACGCGTTAAGTATCCGCTGAGGAGTACGTTCCGA 374  
QY 869 AATTTAAACCTCAAGAAATGACGAGGACCGGACCAAGCGGTGATTAATGATTAAT 928  
DB 373 AATTTAAACCTCAAGAAATGACGAGGACCGGACCAAGCGGTGATTAATGATTAAT 314  
QY 929 TCGATGCAACGCAAAACCTTACCTTACCTTACATGAGGAAATTTCTAGAGATAGA 988  
DB 313 TCGATGCAAAAGGAAAGAACTTACCAAGGCTTGAATGCGGAAATCTCTTGAAGAGA 254  
QY 989 TTAGTG--CTTGGGAAAGCTTAACAGAGTGTGATGCTGTGTCAGTCCGTGCTGA 1047  
DB 253 GGGGTGCTTCCGGGAAACGAGACAGAGTGTGATGCTGTGTCAGTCCGTGCTGA 194  
QY 1048 GATGTTGGGTTAAGTCCGCAACGAGCGCAACCTTGTCAATTAATGGCATTTGCTT 1107  
DB 193 GGTGTTGGGTTAAGTCCGCAACGAGCGCAACCTTGTCAATTAATGGCATTTGCTT 134  
QY 1108 GGGCACTTAATAGACTGCGGTGCAAAACCGAGAAAGTGGGATGAGTCAAGTCC 1167  
DB 133 TGAAGCCCTGAACAGACTGCGGTGCAAAACCGAGAAAGTGGGATGAGTCAAGTCA 74  
QY 1168 TCATGCGCTTAATGAGTGGTGAAGCTTCAACAGTAAATACAAAGCGGTGAGGCCA 1227  
DB 73 TCATGCGCTTAATGAGTGGTGAAGCTTCAACAGTAAATACAAAGCGGTGAGGCCA 14  
QY 1228 ACCCGGAGAGGGG 1240  
DB 13 TCCCGGAGAGGGT 1

RESULT 8  
BH705272 1079 bp DNA linear GSS 20-FEB-2002  
LOCUS BOMKA47TF BO.2.3 KB Brassica oleracea genomic clone BOMKA47,  
DEFINITION genomic survey sequence.  
ACCESSION BH705272  
VERSION BH705272.1 GI:18787746  
KEYWORDS GSS.  
SOURCE Brassica oleracea







## ORIGIN

/clone="BOND28"  
/clone.lib="BO.1.6.2\_KB\_tot"  
/note="Vector: pHO51; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pHO51 using BstXI linkers"

```

Query Match      42.7%; Score 621.8; DB 9; Length 1016;
Best Local Similarity 78.0%; Pred. No. 1.6e-180;
Matches 810; Conservative 0; Mismatches 202; Indels 27; Gaps 4;

OY 200 GGGCATGTCCTATTAAGTCTTGAAGTGGGTAAAGCCCAACGAGCGCATAGT 259
DB 1015 GGGTCGCTGATTAAGTGTAGTGGGTGAGGCAATAGCTTACCAAGCGATGATCAGTACC 956
OY 260 TGGTGTGAGAGACGACGACCACTGGGAGCTGAGACACAGGCGCCAGACTCTTACGGGAG 319
DB 955 TGGTGTGAGAGATGATGACGACCACTGGGAGCTGAGACACGCGCCAGACTCTTACGGGAG 896
OY 320 GCAGCAGTGGGGAATTTTGAACATGGGCGCAAGCTGATCCAGCAATGCCGCTGAGTG 379
DB 895 GCAGCAGTGGGGAATTTTCCG -CAATGGCGCAAGCTGACGAGCAATGCCGCTGAGTG 837
OY 380 AAGAAAGCTTTGGGTTTGAAGCTTTTCACTGAGCAAGAAAGTTACGGTAATAT 439
DB 836 TAGAAGGCTTACGGGCTTGAACCTTTTCCAGAAAGG----- 796
OY 440 CGTGACTGACGCTGATGACGAGAAAGACACCGGCTAACTACGTGCGAGAGCGCGG 499
DB 795 ----ACCAATGACGCTGATGAGGGAATTAAGCATCGGCTAACTGTCGACAGAGCGCGG 740
OY 500 TAAATGATGAGGTGCAAGCGTTAATGGAATTAATGCGGTAAGGCTGACGCGCT 559
DB 739 TAAATGATGAGGTGCAAGCGTTAATGGAATTAATGCGGTAAGGCTGATGAGTGGCT 680
OY 560 TTGTAAGTCAAGTGTAAATCCCGGGCTTAACTTGGGAATTCCTTTAAATACAGG 619
DB 679 TTTTAAGTCCGCGCTCAAAATCCCGGGCTTAACTTGGGAATTCCTTTAAATACAGG 620
OY 620 CTAGAGTGTGCGAGGAGGTGGAATTCATGTGTAGCAGTGAATGCGTATAGATATG 679
DB 619 CTAGAGTGTGCGAGGAGGTGGAATTCATGTGTAGCAGTGAATGCGTATAGATATG 560
OY 680 GAAAGCATGATGAGGAGGAGGAGCTCTGGGTAACTAGCGCTCATGACGAAAGC 739
DB 559 AAGAAACATCAACGAGGAGGAGGAGCTCTGGGTAACTAGCGCTCATGACGAAAGC 500
OY 740 GTGGGAGCAGAAACAGATTAAGATACCTGTGATGCAAGCGCTTAAAGATGTCACTAGT 799
DB 499 TAGGGAGCAGAAACAGATTAAGATACCTGTGATGCAAGCGCTTAAAGATGTCACTAGT 440
OY 800 TGTGGGCTTATTAAGCTTG--GTAACGAAGCTAAGCGTGAAGTTGACCGCTGGGGA 857
DB 439 CGGTGTGCTGATGACCGCTGACAGTGTGATGACGCTTAAAGTATCCCGCTGGGGA 380
OY 858 GTACGCTGCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 917
DB 379 GTACGCTGCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 320
OY 918 TGTGATTAATTCGATGCAACGCGAAACCTTAACCTTGAACATGTAGGGAATTT 977
DB 319 TGTGATTAATTCGATGCAACGCGAAACCTTAACCTTGAACATGTAGGGAATTT 260
OY 978 CTAGAGTATAGTATG--CTTGGGAGCGCTAAGAGGTGCTGATGCTGCTGAGC 1036
DB 259 CTGAAAGAGAGGGGTCTTGGGAGCGGAGACAGGTGCTGATGCTGCTGAGC 200
OY 1037 TCGTGTGATGATGTTGGGTAAAGTCCCGCAAGCAGCGCAACCTTGTATTAATTTGCC 1096
DB 199 TCGTGTGATGATGTTGGGTAAAGTCCCGCAAGCAGCGCAACCTTGTATTAATTTGCC 140
OY 1097 ATCAATTTGTTGGGCTTTATAGACGTGCGGTGCAAAACCGGAGAAAGTGGGATG 1156
DB 139 ACCGTGATGTTGGAAACCTGAAACAGACTCGCGGTGATAGCCGAGAGAAAGTGAAGATG 80

```

```

OY 1157 ACCTCAAGTCTCTCATGGCCCTTATGAGTGGGCTTACACCTAATACAAATGGCGGTACA 1216
DB 79 ACGTCAAGTCTCTCATGGCCCTTATGAGTGGGCTTACACCTAATACAAATGGCGGTACA 20
OY 1217 GAGGTTGCGCAACCGCGCA 1235
DB 19 AAGGTTGCGCAATCCCGCA 1

```

```

RESULT 10
BH656222/c 1031 bp DNA linear GSS 19-FEB-2002
LOCUS BOMJ286TF_BO.2_3_KB Brassica oleracea genomic clone BOMJ286,
DEFINITION genomic survey sequence.
ACCESSION BH656222
VERSION BH656222.1 GI:18714532
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 1031)
Ayle, M., Haas, B.J., Kumar, N., Mu, H., Xiao, Y., Van Aken, S.,
Utecher, T.R., Wortman, J.R., White, O.R., and Town, C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
15805490
COMMENT Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TP
Classes: sheared ends.
Location/Qualifiers
1..1031
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone.lib="BOMJ286"
/clone.lib="BO.2_3_KB"
/note="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHO51 using BstXI linkers"

```

## ORIGIN

```

Query Match      42.4%; Score 617.4; DB 9; Length 1031;
Best Local Similarity 76.5%; Pred. No. 3.8e-179;
Matches 806; Conservative 0; Mismatches 221; Indels 26; Gaps 3;

OY 307 ACTCTTACGGGAGGACGAGTGGGGAATTTTGAACAATGGGCGCAAGCTTGAATCCAGCA 366
DB 1031 ACTCTTACGGGAGGACGAGTGGGGAATTTTCCGCAATGGGCGCAAGCTTGAATCCAGCA 972
OY 367 TGGCGGTGATGATGAAGAAGGCTTGGGTTGTAAAGCTCTTTCAGTGCAGAAAGAAAGT 426
DB 971 TGGCGGTGATGATGAAGAAGGCTTGGGTTGTAAAGCTCTTTCAGTGCAGAAAGAAAGT 918
OY 427 TACGTAATTAATCTGACTCATGACGATTCAGACAGAAAGACACCGGCTAACTACGTG 486
DB 917 -----AGCAATGACGATATCTGGGAAATTAACATCGGCTAACTCTGCTG 875
OY 487 CCAGCAGCGCGGTAAATACGTAAGGTGCAAGCTTATGGAATTAATGCGGCTAAAGG 546
DB 874 CCAGCAGCGCGGTAAATACGTAAGGTGCAAGCTTATGCGGAATGATGCGGCTAAAGG 815
OY 547 TGGCAGCGCGCTTTGATGATGATGGAATCCCGGCGCTTAACTGGGAATTCGCTT 606

```

Db 814 TCTGTAGGTGGCTTTTAAAGTCCGCGCTCAATCCCAAGGGCTCAACCTTGACAGAGCCGT 755  
 Qy 607 TGAATCTACAGGCTTAGAGTGTGGACAGAGGAGGTGGAATTCATGTGTAGAGTGAAT 666  
 Db 754 GGAACCTACCAAGCTTAGTACGATAGGGGCGAGAGGAATTTCCGTGAGCGGTGAAT 695  
 Qy 667 GCGTAGAGATTTGGAAGAACATCGATGGCGAAGGAGCCCTCTGGGTTTAACTGACGT 726  
 Db 694 GCGTAGAGATCGGAAGAACCAACCAAGCGGAAAGCACTTCTGGGCGGACACTACACT 635  
 Qy 727 CATGCACGAAGCGTGGGAGGCAACAGGATTAGATCCCTGTGTAGTCACGCCCTTAAC 786  
 Db 634 GAGAACCAAGAGCTAGGGGAGGGAATGGGATTAGATACCCAGTGTCTACCGCTTAAC 575  
 Qy 787 GATGTCAACTAGTGTGTGGGCTTATTAGGCTTG--GTACGAACTTAACGGTGAAGT 844  
 Db 574 GATGATCTAGGCGCTGTGCTATCGACCCCTGAGTGTCTAGCTAAACGCTTAAGTA 515  
 Qy 845 GACCGGCTGGGAGTACGGGTGCGAAGTTAAACTCAAGGAATTGACGGGGACCGGAC 904  
 Db 514 TCCCGCTGGGAGTACGTTGCGAAGAAATGAATCTCAAGGAATTGACGGGGCGGAC 455  
 Qy 905 AAGCGGTGATTTATGATTAATTCATGCAACCGCAAAAACCTTACCTACCTTGAACA 964  
 Db 454 AAGCGGTGAGCATGTGTTTAATTCGATGCAAGCAAGAACCTTACAGGGCTTGAACA 395  
 Qy 965 TGTACGGAATTTTCTAGAGATGATTAGTG-CTTGGGGAACCTTAACAGGTGCTCAT 1023  
 Db 394 TGCCGCGAATCTCTTGAAGAGAGGGGTGCTTGGGGAACCGGACACAGGTGTGAT 335  
 Qy 1024 GCGGTGCTGACGTGTGTGCGAGATGTGGGTTAAGTCCCGCAACGAGCGCAACCTT 1083  
 Db 334 GCGGTGCTGACGTGTGTGCGAGATGTGGGTTAAGTCCCGCAACGAGCGCAACCTT 275  
 Qy 1084 GTCAATTAATTCATCATTTGTTGGGCACTTAAATGAGACTGCCGTGACAAACCGGAG 1143  
 Db 274 GTGTTTATTTGCAACCGTTGAGTTTGAACCTGAACAGACTGCCGTGATAGCGGAG 215  
 Qy 1144 GAAGGTGGGAGTACGTCAAGTCTCATGAGCCCTTATGGGTAGGCTTCAACGTAATAC 1203  
 Db 214 GAAGGTGAGATGACGTCAAGTCAATGATCCCTTATGCGCGGCGACACGCTGATAC 155  
 Qy 1204 AATGCGGTGACAGAGGTTGCGCAACCGGAGGGGAGCTAATTCAGAAAGCGCGTGC 1263  
 Db 154 AATGCGGTGACAGAGGTTGCGCAACCGGAGGGGAGCTAATTCAGAAAGCGCGTGC 95  
 Qy 1264 TAGTCCGATCGAGTCTGCAACTGCACTCCGTGAAGTCGGAATCGTATATCGCGGA 1323  
 Db 94 CAGTTCCGATTTGCAAGCTGCAACTGCGCTGCATGAAGCGGAAATGATATATGCGCG 35  
 Qy 1324 TCAAGCATGTCCGGGTGAATACGTTCCCGGGTCT 1356  
 Db 34 TCAGCCATACGCGGGAATTCGTTCCCGGGCT 2

RESULT 11  
 B2450751 1077 bp DNA linear GSS 13-DEC-2002  
 LOCUS BOMB063TR.B0.1.6.2 KB tot Brassica oleracea genomic clone BOMB063,  
 DEFINITION genomic survey sequence.  
 ACCESSION B2450751  
 VERSION B2450751.1 GI:26721334  
 KEYWORDS GSS.  
 ORGANISM Brassica oleracea  
 Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 1077)  
 Ayale, M., Haas, B., J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
 Utecherback, T. R., Wortman, J. R., White, O. R. and Town, C. D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis

JOURNAL  
 PUBMED  
 15805490  
 COMMENT  
 Other\_GSSs: BOMB063TR  
 Contact: Chris Town  
 TIR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Classes: sheared ends.  
 FEATURES  
 source  
 1..1077  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="T0100DH3"  
 /db\_xref="taxon:3712"  
 /clone\_lib="BOMB063"  
 /clone\_lib="B0.1.6.2 KB tot"  
 /note="Vector: pHD01; Site: 1: BexXI, 1.6-2 kb sheared  
 total DNA inserted into pHD01 using BexXI linkers"

ORIGIN  
 Query Match 42.2% Score 615.2; DB 9; Length 1077;  
 Best Local Similarity 77.5%; Pred. No. 1.9e-178;  
 Matches 783; Conservative 0; Mismatches 223; Indels 4; Gaps 3;

Qy 448 ATGACGGTATGACAGAGAAGACACCGGCTTAACCTAGCTGCGACAGCGCGGTAAATAGT 507  
 Db 68 ATGACGGTATGCGGAGATAGATCGGCTAACTGTGCGACAGCGCGGTAAATAGT 127  
 Qy 508 AGGTCGCAAGGCTTAATTCGAATTAATCTGCGGTAAAGGTGCGAGCGGCTTGTAGT 567  
 Db 128 AGGATGCAAGGCTTAATTCGAATTAATCTGCGGTAAAGGTGCGGTGCTTTTAAAT 187  
 Qy 568 CAGATGTAAGTATCCCGGGCTTAACCTGGGAAATTCGTTGAATCTCAAGGCTAGAGTG 627  
 Db 188 CCGCGGTCAATCCGAGGCTCAACCTGACAGAGCGGTGAAACTTACAGCTTGAAT 247  
 Qy 628 TGGCAGAGGAGTGAATTCATGTGTAGCAGTGAATATGCGATATGAGAACA 687  
 Db 248 CCGTAGGGGAGAGGAAATTTCCGCTGAGAGCGGTGAATGTGATGATCGAAGAACA 307  
 Qy 688 TCGATGCGAAGGACGCTCTGCGGTTAACCTACGCTCATGACGAAAGCGTGGGAG 747  
 Db 308 CCAACGCGAAGGACCTCTGCTGGCCGACACTGACACTGAGAACCAAGCTGAGGAG 367  
 Qy 748 CAACAGGATTAAGTATCCGTGTAGTCCAGCGCCCTTAACGATGCACTAGTTTGGGC 807  
 Db 368 GGAATGGATTAAGTATCCCGATGCTCTAGCCGTAAACATGATACTAGGCGCTGTGC 427  
 Qy 808 CTATTAGGCTTG--GTACGAACTTAACCGGTGAAGTTGACCGCTGGGGAGTACGCTC 865  
 Db 428 GTATGCAACCGGTGACGTGTAGCTTAACCGGTGAAGTATCCGCTGGGAGTACGCTC 487  
 Qy 866 GCAGATTAATACTCAAGGAATTTGACGGGACCCCGCAACAGCGGTGATATGATGATT 925  
 Db 488 GCAAGATTAATACTCAAGGAATTTGACGGGACCCCGCAACAGCGGTGATATGATGATT 547  
 Qy 926 AATTGATGCAAGCGGAAAACTTACCTACCTTGAATGATGCAATTTTCTAGAGAT 985  
 Db 548 AATTGATGCAAGCGGAAAACTTACCTACCTTGAATGATGCAATTTTCTAGAGAT 607  
 Qy 986 AGATTAGTG-CTTGGGAGCTTAACACAGGTGTGATGAGCTGTCTGACGCTGTGTC 1044  
 Db 608 AGAGGGGTGCTTGGGAGCGGACACAGGTGTGATGAGCTGTCTGACGCTGTGTC 667  
 Qy 1045 TGAATGTTGGGTAAAGTCCCGCAGACAGGCAACCTTGTCAATTAATGCAATTTG 1104  
 Db 668 TAAAGTGTGGGTAAAGTCCCGCAGACAGGCAACCTTGTCAATTAATGCAATTTG 727  
 Qy 1105 GTTGGGCACTTTATGAGACTGCCGCTGACAAACCGGAGGAAGGTGGGATGACGTCAAG 1164

Accession	Sequence	Position
Db	GTTTGAACCCCTGAACACACTGCGCGTATTAACCGAGGAAGGTGAAGATGAAGCTCAG	787
Oy	TTCTCATGAGCCCTTATGGGTAGGGCTTTCACAGTAAACAATGGCGCTACAGAGGTTG	1224
Db	TCATCATGTCCTTATATGCGCCGTGGGGAACACAGTGTCTACAATGGCCGGACAAAGGGCTCG	847
Oy	CCAACCCGCGAGGGGGAAGCTAATCTCAGAAAAGCGGTGTAGTCGGATCGGAAGTCTGCA	1284
Db	CGATCCCGCGAGGGGAGACTTAATCTCAAAAACCCGCTCTCAGTTGGATATGCAAGCTTCA	907
Oy	ACTCGACTCCGTAAGTGGAAATCGCTATATCGCGGATCAG-CATGTCCGGTGAATA	1343
Db	ACTCGCTCTGACGAAGCCGGAATGCTATGTAATGCGCGTACGCCATACGGGGGTGAATT	967
Oy	CGTTCCGGGCTCTTGTAAACACACCGCCGCTCACATATGGAGAGCTGGCCATGCCCAAAAGT	1403
Db	CGTTCCCGGGCTTGTAAACACACCGCCGCTCACATATGGAGAGCTGGCCATGCCCAAAAGT	1027
Oy	GGTAGCTTAACCGTAAAGGAGGGCGCTTGCGACGGTGAATTCATAGTCTG	1453
Db	GGTAGCTTAACCGCGAAGAGGGGGGTGCGGAAGGACAGGGCTAGTACTG	1077

RESULT	12
BZ463550/c	
LOCUS	BZ463550 1044 bp DNA linear GSS_13-DEC-2007
DEFINITION	BONKCBETP BO_1.6_2_KB tot Brassica oleracea genomic clone BONKC96,
	genomic survey sequence.
ACCESSION	BZ463550
VERSION	BZ463550.1 GI:26749465

SOURCE	ORGANISM
Brassica oleracea	Brassica oleracea

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosoids; eurosid II; Brassicales; Brassicaceae; Brassica.

REFERENCE  
1. (BASES 1 CO 1074)  
Ayele, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,

**TITLE** Whole genome shotgun sequencing of *Brassica oleracea* and its application to gene discovery and annotation in *Arabidopsis*

JOURNAL OF CLIMATE 20 (1997) 1201-1208  
 PUBMED 15805490  
 0894-8755(199707)20:7:1-0

Contact: Chris Town  
TIGR

**Tel: 301-838-3523**  
**Fax: 301-838-0208**

Email: [catowin@rci.org](mailto:catowin@rci.org)  
DNA is from a doubled haploid provided by Tom Osborn.

**FEATURES**  
**Class:** sheared ends.  
**Location/Qualifiers**

/organism="Brassica oleracea"  
/col="Broccoli"

```
/strain="T01000DH3"
/db_xref="taxon:3712"
```

```
/clone lib="BO_1.6_2_KB_tot"
/notes="Vector: pHOSt1; Site 1: BstXI; 1.6-2 kb sheared"
```

ORIGIN	COLL	DATE	TIME	FILE	NAME	TYPE	STATUS	REMARKS
COLL	DATE	TIME	FILE	NAME	TYPE	STATUS	REMARKS	

Query Match	42.1%;	Score 613.2;	DB 9;	Length 1044;
Best Local Similarity	78.5%;	Pred. No. 7.6e-178;		

448 ATGAGCGCTTCCACACAGAGACACCGGCCTAACTACGTGCCAGCAGCGCGGTAAATTCGT 507

Db 1000 ATGACGTATCTGGGAATAAGCATCGGCTAACTCTGTGCACAGCAGCCCGGTAATACAG 941

QY	508	AGGCTGCMAACGTTTAATCGGAATTTACTGGCCGTTAAAGGGTCCGACGGCGCTTTGTAGT	567
Db	940	AGGATGCACGCTTATCCGGAAATGATGGCCGTTAAACGCTCTGTAGGTTGGCTTTTGTAGT	881
QY	568	CAGATGTGAAATCCCCGGGCTTAACTCTGGGAATTTGCGTTGAAACTCAAGGCTAGAGTG	627
Db	880	CCGCGCTGCAATCCCAAGGGCTCAACCCCTGGACAGCCGGTGGAAACTCAACAGCTTGAAGTA	821
QY	628	TGGCAGAGGGAGGTGGTAATTCATGTGTAGCAGTGAATGGGTAGAGTATGGAAGAACA	687
Db	820	CGGTAGGGGCGACAGAGGAATTTCCGGTGGACCGGTGAATGTGTAGATCGGAAAGAACA	761
QY	688	TCGATGCGCAAGGACGAGCCTCCTGGGTTAACTAGAGCTCATGCAAGAAACGTTGGGAG	747
Db	760	CCAACGGCGGAAGGACACTGTGCTGGCCGGAACCTGACACTGAGAGACAAAGACTAGGGGAG	701
QY	748	CAAAACAGGATTAGATCCCTGTGTATGTCACGCCCTTAAACGATGTTCMACTAATTGTGGGC	807
Db	700	CGAATGGGAATTAGATACCCCACTAGTGTCTTACGCCCTTAAACGATGTACTAGGCCCTGTGC	641
QY	808	CTTATTATAGGCTTG--GTAAACAACCTAACCGGTAAAGTGAACCGCTGGGAGTAGCGCTC	865
Db	640	GTATTCGACCCCTGCAGAGCTGTGTACTTAACGGGTTTAAGTATCCCGCTGGGAGATACGCTTC	581

866 GCAAGATTAAACTCAAAGGATTGACGGGGACCCGCACAAGCGGTGATTTATGTGGATT 925

[illegible]

520 AATTGATGCAAGCGAAGACTTACCAGGGCTTACATGCCGGAATCCTTTGAAG 461

986 AGATTACTG-CTTCGGAGCCTAACACAGGTGTCATGGCTGTCTAGCTCTGTTCG 104

Db 460 AGAGCGGTCCTTCGGAAACGGACACAGAGTGTGCAAGCAGTCGACAGCAGTCCCG 102

400 TAACGCTTGGAGTTAAGTCCCGCAACGAGCGCAACCTTCGTGTTAGTTCGCCACCGTTGA 341

1105 GTTGGCACTTAAATGAGACTGCGGTGACAAACCGAGGAAGTGGGATGACGTCAG 116

Db 340 GTTGAACCTGAACAGACTGCCGCTGATTAAGCCGAGGAGGTGAGGATGACCTCAG 261

[illegible]

1225 CCAACCCGCGAGGGGAGCTAATCTAGAAAGCGGTCGTAGTCCGGATCGGACTGTGCA 128

Db 220 CGATCCCGCGAGGGTGAGCTAACTCCAAAAACCCGTCCTCAGTTCGATTGCAAGGCTGCA 161

101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534  
 535  
 536  
 537  
 538  
 539  
 540  
 541  
 542  
 543  
 544  
 545  
 546  
 547  
 548  
 549  
 550  
 551  
 552  
 553  
 554  
 555  
 556  
 557  
 558  
 559  
 560  
 561  
 562  
 563  
 564  
 565  
 566  
 567  
 568  
 569  
 570  
 571  
 572  
 573  
 574  
 575  
 576  
 577  
 578  
 579  
 580  
 581  
 582  
 583  
 584  
 585  
 586  
 587  
 588  
 589  
 590  
 591  
 592  
 593  
 594  
 595  
 596  
 597  
 598  
 599  
 600  
 601  
 602  
 603  
 604  
 605  
 606  
 607  
 608  
 609  
 610  
 611  
 612

1344 CGTTCGCCGGTCTTGTACACACCGCCCGTCACACCATGGGAGTGGTTTCACCGAAGCA 1400

Db  
100 CGTTCGCCGGCTTGTACACACCGCGGTCACTATGGAGCTGGCCATGCCCCGAACTC 41

QY 1404 GGIAGICIMACCGTAAAGAGGCG 1427

BZ494182/C  
KBSVDL 13  
07/06/1993

DND 1061 bn 11near GSS 16-DEC-2000

**DEFINITION** `BONHR54TR_BO_1.6.2_KB_FOT` Brabbsica oriented genomic clone bonhr54tr genomic survey sequence.

VERSION	B2494182.1	GI:27004681
---------	------------	-------------

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustroids II; Brassicales; Brassicaceae; Brassica.

AUTHORS Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Uteyback, T.R., Wortman, J.R., White, C.R. and Town, C.D.

TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis

JOURNAL Genome Res. 15 (4), 487-495 (2005)

PUBMED 15805490

COMMENT Other\_GSSs: BONHR54TF  
Contact: Chris Town  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source  
1..1051  
Location/Qualifiers  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="T0100DH3"  
/db\_xref="taxon:3712"  
/clone="BONHR54"  
/note="Vector: pHD1; Site: 1; BcXI: 1.6-2 kb sheared  
total DNA inserted into pHD1 using BcXI linkers"

ORIGIN

Query Match 42.1%; Score 613.2; DB 9; Length 1051;  
Best Local Similarity 78.5%; Pred. No. 7,66-178;  
Matches 772; Conservative 0; Mismatches 208; Indels 4; Gaps 3;

448 ATGACGGTATCGACAGAAAGACACCGGCTAACTACGTGCGACGACCGCGGTAAATACGT 507  
1000 ATGACGGTATCGGGAATTAAGCATCGGCTAACTGTGCGACGACCGCGGTAAATACAG 941

508 AGGATGCAAGCTTAATCGGAATTAATCGGCTAAAGGATGGCGAGGGCTTTGTAAGT 567  
940 AGGATGCAAGCGTTATCCGGAATGATTTGGCGTAAAGCGTGTAGTGCGCTTTTAAGT 881

568 CAGATGTAATATCCCGGCTTTAACTCGGGAATTCGTTTGAATACCAAGGCTAGAGTG 627  
880 CCGCGCTAAATCCAGGGCTCAACCCCGACAGCGCGGTGAAACTACCAAGCTTGAATG 821

628 TGGCAGAGGAGGTGAATTCATGTGTAGCAGTGAATGCGTAGAGATATGGAAGACA 687  
820 CGTAGGGGCGAGAGGAATTTCCGGTGAAGCGGTGAATGCGTAGAGATCGGAAGACA 761

688 TCGATGGGAGGAGCGCTCTCGGTTAACTGACGTGACGCAAGAACCGTGGGAG 747  
760 CCAACGGGGAAGCACTGCTGGGCGCAACACTGAGAGAGCAAGACTAGAGGAG 701

748 CAACAGAGTATAGTACCTGTAAGTCCAGCGCTTAAAGATGTAAGTCACTAGTGTGGGC 807  
700 GCAATGGGATTAAGTATCCCGAGTAGTCTTAGCCGTTAAAGATGATATCTAGCGCTGTGC 641

808 CTATTAAGGCTTG--GTAAAGAGCTAAACGGGTGAAGTGAACCGCTTGGGAGTACGGTGC 865  
640 GTATCGAACCGGTGCGTCTGTAGCTAACCGGTTAAAGTATCCCGCTGGGAGTACGTTTC 581

866 GCAAGATTAATACTCAAGGAATTAAGCGGGAACCGGCAAGCGGTGATTAATGTGGATT 925  
580 GCAAGAAATGAATCAAGGAATTAAGCGGGAACCGGCAAGCGGTGAGGATGTGGATT 521

926 AATTCGATGCAAGCGGAATACTTACCTTACCTGATGATGAGGATTTTCTAGAGAT 985

Db 520 AATTCGATGCAAGCGGAATACTTACCTTACCGGGCTTGAATGATCCGCGAATCTTTGAAG 461

Qy 986 AGATTATG-CTTGGGGAACCTTAACACAGGTGCTGATGGCTGTCTGACGCTGTGTCG 1044

Db 460 AGAGGGGCTTTCGGGAACCGGACACAGGTGTGATGCTGTCTGACGCTGTGTCG 401

Qy 1045 TGAATGTGGGTTAAGTCCCGCAAGCGCAACCTTGCATTAATTTGCGATCTTGG 1104

Db 400 TAAAGTGTGGTTAAGTCCCGCAAGCGCAACCTTGCATTAATTTGCGATCTTGG 341

Qy 1105 GTTGGGCACTTTAATGAGACTGCGGTGACAAACCGGAGGAGTGGGATGAGCTCAAG 1164

Db 340 GTTTGGAACCTTGAACAGACTGCGGTGATAGCGCGGAGGAGTGAAGATGAGTCAAG 281

Qy 1165 TCTCATGAGCCCTTAATGAGGTGAGGCTTCAACGTTAATCAATGCGCGTACAGAGGTTG 1224

Db 280 TCATCATGAGCCCTTAATGAGGTGAGGCTTCAACGTTAATCAATGCGCGGAGGAGTGC 221

Qy 1225 CCAACCGCGAGGGGAGGCTAATCTCAGAAAGCGCGTGTAGTCCGAGTCCGAGTTCGA 1284

Db 220 CGATCCGCGAGGGTGAAGCTAATCTCAGAAAGCGCGTCTTCAATGAGTTCAGAGCTGCA 161

Qy 1285 ACTGACTCCGTGAAGTGTGAGATGCTAATGATGCGGATGAG-CATGTCCGCTGAATA 1343

Db 160 ACTGCGCTGATGAAGCGGAATCGTAGATGCGCGGTGAGCATACGCGGTGAATT 101

Qy 1344 GCTTCCCGGCTTGTGACACACCGCGCTCACACATGAGGAGTGGTTTACCGAAGCA 1403

Db 100 CGTTCGCGGCTTGTGACACACCGCGCTCACACTATGAGAGTGGCATGCGCAAAAGTC 41

Qy 1404 GGTAGTCTAACCGTGAAGAGGGCG 1427

Db 40 GTTACCTTAACCGCAAGAGGGCG 17

RESULT 14  
CK280527 889 bp mRNA linear EST 02-AUG-2004  
LOCUS  
DEFINITION  
CK280527.1 GI:39850196  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TIGR  
JOURNAL  
COMMENT  
Other ESTs: EST743250  
Contact: Robin Buel  
The Institute for Genomic Research  
9712 Medical Center Dr. Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES  
source  
1..889  
Location/Qualifiers  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NEM117"  
/tissue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_id="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"

/note="Vector: pCMVSPORT6.1; Site\_1: EcoRI; Site\_2: NotI; supplier: RNA was isolated from *Nicotiana benthamiana* tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (*Pseudomonas syringae* pv tomato 12 hr; *Xanthomonas campestris* pv *campestris* 12 hr, 18hr; *Pseudomonas syringae* pv *phaseolicola* 18hr, and *Xanthomonas campestris* pv *vesicatoria* 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

## ORIGIN

```

Query Match      41.3%; Score 601.4; DB 7; Length 889;
Best Local Similarity 81.4%; Pred. No. 3.3e-174;
Matches 709; Conservative 0; Mismatches 161; Indels 1; Gaps 1;

1  ATTGAACGCTGGCGGCGAGCTTTTACATGCAAGTGAACGCGACGACGCGATGCTTGCA 60
18  ATTGAACGCTGGCGGCGAGCTTTTACATGCAAGTGAACGCGACGACGCGATGCTTGCA 77
61  CTGGTGGCGAGTGGCGGAGCGGGAGTATGCAATCGGAACGATCCAGAAAGAGGGGGATA 120
78  CTGGTGGCGAGTGGCGGAGCGGGAGTATGCAATCGGAATCTGCTGTGATGGGGGATA 137
121  ACGCATCGAAAGATGTCTAATACCGCATATATCTTAAGAGAGAAAGACGAGGATCGAAA 180
138  ACGCTCGGAAAGCGAGCTTAATACCGCATATCGTCTTAAGAGAGAAAGACGAGGATCG 197
181  GACCTTGGCTTTTGGAGCGCGGATGTCTGATTAAGTGTGGTGAAGGCTTAC 240
198  GGCCTTGGCTTACATGATGAGCTGATGCGATTAAGTGTGATGAAGGCTTAC 257
241  CAAGGCGAAGTACATGATGTCTGATGAGGAGACGACGACGACGACGACGACGACG 300
258  CAAGGCGAAGTACATGATGTCTGATGAGGAGACGACGACGACGACGACGACGACG 317
301  GCCCACTCTTACGCGGAGGAGCAAGTGGGGAATTTTGGACATAGGCGGACGCTGATC 360
318  GTCCAGACTCTTACGCGGAGGAGCAAGTGGGGAATTTTGGACATAGGCGGACGCTGATC 377
361  CAGCAATGCGCGCTGATGAGTGAAGAGCGCTTGGGTTGAAGCTCTTTCAGTGAAGA 420
378  CAGCCTGCGCGCTGATGAGTGAAGAGCGCTTGGGTTGAAGCTTTCAGTGAAGA 437
421  AAGGTTACGTTAATATCTGATGATGATGATGATGATGATGATGATGATGATGATG 480
438  AAGGTTACGTTAATATCTGATGATGATGATGATGATGATGATGATGATGATGATG 497
481  TACGTGCGACGACGCGGATATACGTAAGGTCAGACGCTTAATCGGAATTACTGGGCGT 540
498  TCTGTGCGACGACGCGGATATACGTAAGGTCAGACGCTTAATCGGAATTACTGGGCGT 557
541  AAGGTTGCGCGAGCGGCTTTTGAAGTCAAGTGAAGTCCCGGGCTTAACCTGGGAT 600
558  AAGGCGGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 617
601  TGGCTTTGAATCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 660
618  TGCATCCAAAACCTGGCAAGCTAGATGATGATGATGATGATGATGATGATGATGATG 677
661  TGAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
678  TGAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737
721  GACGCTATGCAAGAAAGCGTGGGAGCAACAGGATTAATACCTGTTGATGATGATG 780
738  GACGCTATGCAAGAAAGCGTGGGAGCAACAGGATTAATACCTGTTGATGATGATG 797
781  CTAAACGATGCACTAGTTGTTGGGCTTAATAG -GCTTGTGAACGAAGTAAAGCGG 839
798  GTAAACGATGCACTAGTTGTTGGGAGCGCTTGAAGCTTGAAGTGGCGACGTAAGCA 857

```

QY 840 AAGTTGACCGCTGGGAGTACGATCGCAAG 870  
 DB 858 AAGTTGACCGCTGGGAGTACGATCGCAAG 888

RESULT 15  
 LOCUS CK280528  
 DEFINITION EST743250 *Nicotiana benthamiana* mixed tissue cDNA library,  
 normalized, full-length *Nicotiana benthamiana* cDNA clone NBNA117 5'  
 end, mRNA sequence.

ACCESSION CK280528  
 VERSION CK280528.1 GI:39850200

KEYWORDS EST  
 SOURCE *Nicotiana benthamiana*  
 ORGANISM *Nicotiana benthamiana*

REFERENCE 1 (bases 1 to 884)  
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
 Staekwitz, B., Jin, H., and Baker, B.

TITLE Generation of EST sequences from *Nicotiana benthamiana*  
 JOURNAL Unpublished (2003)

COMMENT Other ESTs: EST743249

Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org  
 Clones can be requested from the University of Arizona Genomics  
 Institute via <http://genome.arizona.edu/orders/>.  
 Seq primer: CAG GAA ACA GCT ATG ACC.

## FEATURES

## source

1..884  
 /organism="Nicotiana benthamiana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4100"  
 /clone="NBNA117"  
 /tissue\_type="abiotic and biotic stress-treated leaves,  
 callus tissue and root tissue"  
 /lab\_host="DH10B-TonA"  
 /clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
 library, normalized, full-length"  
 /note="Vector: pCMVSPORT6.1; Site\_1: EcoRI; Site\_2: NotI;  
 supplier: RNA was isolated from *Nicotiana benthamiana*  
 tissues that include callus, roots from liquid culture  
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen  
 challenged leaves (*Pseudomonas syringae* pv tomato 12 hr;  
*Xanthomonas campestris* pv *campestris* 12 hr, 18hr;  
*Pseudomonas syringae* pv *phaseolicola* 18hr, and *Xanthomonas*  
*campestris* pv *vesicatoria* 18hr). RNA was isolated from  
 these tissues and pooled in approximately equal molar  
 amounts."

## ORIGIN

```

Query Match      41.0%; Score 597.4; DB 7; Length 884;
Best Local Similarity 81.3%; Pred. No. 5.7e-173;
Matches 705; Conservative 0; Mismatches 161; Indels 1; Gaps 1;

1  ATTGAACGCTGGCGGCGAGCTTTTACATGCAAGTGAACGCGACGACGCGATGCTTGCA 60
18  ATTGAACGCTGGCGGCGAGCTTTTACATGCAAGTGAACGCGACGCGATGCTTGCA 77
61  CTGGTGGCGAGTGGCGGAGCGGGAGTATGCAATCGGAACGATCCAGAAAGAGGGGGATA 120
78  CTGGTGGCGAGTGGCGGAGCGGGAGTATGCAATCGGAATCTGCTGTGATGGGGGATA 137
121  ACGCATCGAAAGATGTCTAATACCGCATATATCTTAAGAGAGAAAGACGAGGATCGAAA 180
138  ACGCTCGGAAAGCGAGCTTAATACCGCATATCGTCTTAAGAGAGAAAGACGAGGATCG 197
181  GACCTTGGCGCTTTTGGAGCGCGGATGTCTGATTAAGTGTGGGATGAAGGCTTAC 240

```

```

Db      198 GGCCTTGGCTATCAGATGAGAGCTTGTGATTTAGCTAGTTGTGAGTATGGCTCAC 257
Qy      241 CAAGCGCAGATCAGTACTGTTGTCTGAAGAGACGACCGCACTGGAGTGAAGACAG 300
Db      258 CAAGCGCAGATCCTTAATCTGCTGAGAGATGATCACTGCACTGGAACTGAGACAG 317
Qy      301 GCCCAGACTCTTACGCGGAGGCGAGCTGGGGAAATTTGGACAAATGGGGCGAAAGCTGATC 360
Db      318 GTCCAGACTCTTACGCGGAGGCGAGCTGGGGAAATTTGGACAAATGGGGCGAAAGCTGATC 377
Qy      361 CAGCAATCCCGCGTGTGAGAGAGAGCCCTCGGGTGTGTAAGCTCTTTCACTGAGAGAGA 420
Db      378 CAGCATATCCCGCTGTGTGAGAGAGTCTTCGGATTGTGAACACTTTAAGTTGGAGAGA 437
Qy      421 AAGGTTACGGTAATTAATCTGACTCATGACGCTATGACAGAAAGAACCGGCTAAC 480
Db      438 AGGCGAGTTACCTTAATCAGTATGTTGTTGACGTTACGACAGAAATTAAGCACCGGCTAAC 497
Qy      481 TACGTGCGACGACCGCGGTTATAGTGAAGGTGCAAGGCTTAATCGGAATTACTGGGGGT 540
Db      498 TCTGTGCGACGACCGCGGTTATACAGAGGTGCAAGGCTTAATCGGAATTACTGGGGGT 557
Qy      541 AAAGGTCGCGACGCGCTTTGTGATGATGATAATCCCGGGCTTAACTGGGAT 600
Db      558 AAAGGCGCGCTGATGCTTTGTGATGATGATAATCCCGGGCTTAACTGGGAT 617
Qy      601 TCGCTTGAATCTACAGAGCTAGAGTGTGCGACAGAGGAGTGAATTCATGTGACAG 660
Db      618 TGCATCCAAATCTGCGCAAGCTAGAGTATGTGAGAGGAGTGTGGAATTTCTGTGACGG 677
Qy      661 TGAATGTGTAAGATATGAGAAACATCGATGGGAGGAGCGCTCTCGGGTTAACT 720
Db      678 TGAATGTGTAAGATATGAGAAACATCGATGGGAGGAGCGCACTGAGCTATACT 737
Qy      721 GACGCTCATGACGAAACGCTGGGGAGCAAAACAGATTAGATACCTGTGATCCAGCC 780
Db      738 GACACTGAGTGTGCGAAACGCTGGGGAGCAAAACAGATTAGATACCTGTGATCCAGCC 797
Qy      781 CTAAACGATGTCAACTGATTTGTGGGCTTATTAAG-GCTTGTGAACGAAGCTAACGCGTG 839
Db      798 GTAACGATGTCAACTGATTTGTGGAGCCTTGAAGCTTATGATGCGCAGCTAACGCAAT 857
Qy      840 AAGTTGACCGCTGGGGAGTACGCTG 866
Db      858 AAGTTGACCGCTGGGGAGTACGCGC 884

```

RESULT 16  
BH647750/c 1054 bp DNA linear GSS 19-FEB-2002  
LOCUS BOMND31TR BO\_2\_3 KB Brassica oleracea genomic clone BOMND31,  
DEFINITION genomic survey sequence.  
ACCESSION BH647750  
VERSION BH647750.1 GI:18705298  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 1054)  
Ayele,M., Haas,B.,J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Uteyback,T.R., Wortman,J.R., White,O.R. and Town,C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
JOURNAL PUBMED 15805490  
COMMENT Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208

```

FEATURES
    source
        1..1054
            /organism="Brassica oleracea"
            /mol_type="Genomic DNA"
            /strain="TO1000DH3"
            /db_xref="taxon:3712"
            /clone="BOMND31"
            /note="Vector: pHD31; Site 1: BstXI; 2-3 kb sheared
            genomic DNA inserted into pHD31 using BstXI linkers"
ORIGIN
    Query Match 40.9%; Score 596.2; DB 9; Length 1054;
    Best Local Similarity 77.6%; Pred. No. 1.4e-172;
    Matches 772; Conservative 0; Mismatches 218; Indels 5; Gaps 4;
Qy      467 AAGCACCCTTAATCTAGCTGCGACAGCCGGCTTAATCGTAGGGTGCAGACCTTAATCG 526
Db      1054 AAGCATCGGCTTAATCTGTGCGACAGCGCGGTAAATACAGAGATGCAGACCTTAATCG 995
Qy      527 GAATTAATCTGGCGTAAAGGAGTGCAGCGCGCTTTGTATGATGATGATAATCCCGGG 586
Db      994 GAATGATTTGGCGTAAAGCGTGTAGTGGCTTT-TAAGTCCCGCTTAATCCAGGG 936
Qy      587 CTTAACCTGGGAATTCGTTTGAATCTAACAGGCTAAGTGTGCGAAGGAGGTGAT 646
Db      935 CTCAACCTGGACAGCGCGGTGAAATCTAACAGCTTGATGATCGGATGAGGCGAGGGAT 876
Qy      647 TCCATGTGTACACTGAAATGCGGTAGAGATGTGAAAGAACTGATGGCGGAAGGAGCT 706
Db      875 TTCCGTGAGACGGTGAATGCGGTAGAGATGAAAGCAACCAACGCGCAACCACTCT 816
Qy      707 CCTGGGTTAACTAGTACGCTCATGACGAAAGCGTGGGGAGCAAAACAGATTAGATACC 766
Db      815 GCTGGGCGGACACTGACACTGAGAGACGAAAGCTAAGGGAGCAATGGGATTAATATACC 756
Qy      767 TGTGATGCCAGCGCCCTTAACAGATGTCAATGTTGTGGGCTTATTAAGCTTG-GTAA 824
Db      755 CAGTAGTCTTAAGCGGTAAACGATGATATCTAGGCGCTGTGGTATCAACCGCTCAGTGC 696
Qy      825 CGAAGCTTAACCGCTGAAGATTGACCGCTGGGGAGTATGCGTGAAGATTAACTCAAG 884
Db      695 TGTGCTTAACCGCTTAAGTATCCGCTGGGGAGTATGCGTGAAGATTAACTCAAG 636
Qy      885 GAATTAACGGGAGCCCGACAAAGCGGTGATATTATGTGATTAATTCATGACGCGAAA 944
Db      635 GAATTAACGGGAGCCCGACAAAGCGGTGAGATGTGTTATTCATGCAAGCGAAG 576
Qy      945 AACCTTAACCTTGAACATGTAGCGAATTTCTAGAGATGATTAAGTGTG-CTTCGAGAA 1003
Db      575 AACCTTAACAGGCTTGAACATGTAGCGGCAATCTCTTGAAGAGAGGGGTGCTTCGAGAA 516
Qy      1004 CGCTTAACAGAGTCTGATGAGCTGTGTGAGCTGTGTGTGAGATTTGGGTTAAGTGC 1063
Db      515 CGCGAACAAGTGTGATGTGCTGTGTGAGCTGTGTGAGGTTGGGTTAAGTGC 456
Qy      1064 CCGCAACGAGGCAACCTTGTCAATTAATGSCATCAATTTGTGGGCACTTTAATGAGA 1123
Db      455 CCGCAACGAGGCAACCTTGTGTATTAATGTTATGCCAGCTTGAATTTGAACCTTAACAGA 396
Qy      1124 CTGCGGTGACAAACCGAGAGAGTGGGATGACGTCAAGTCTCAATGCGCTTAATGGG 1183
Db      395 CTGCGGTGATTAACCGAGAGAGTGAAGATGACGTCAAGTCAATGCGCTTAATGGC 336
Qy      1184 TAGGGCTTACAGCTAATTAATGAGCGCGTACAGAGGTTGCCAACCCGAGAGGGAGC 1243
Db      335 CTGGGCGACACACGTGTATCAATGCGCGGACAAAGGTTCCGATCCCGCAGAGGTAGAG 276
Qy      1244 TAATCTGAAGAAAGCGGTCGTAGTCCGGATGAGATGCAACTCGAATCTCGTGAAGTGC 1303

```

```

Db      275  TAACTCCAAAAACCGCTCCTCAGTTCGAGATTCAGAGCTGCAACTGCGCTGCATGAAAGCG 216
Qy      1304 GAATGCTAGTAATGCGCGGATAG-CATGTGCGGGGAAATAGCTTCCGGGCTTGTACA 1362
Db      215  GAATGCTAGTAATGCGCGGATAGCGGCGGATAGCTTCCGGGCTTGTACA 156
Qy      1363 CACCGCCGTCACACCATGAGAGTGGTTCACCAAGACAGTAGTCTAACGTAAGA 1422
Db      155  CACCCCGCTGACATATGAGGAGCTGGCCATGCGCAAGTCTTAACTTAACTGCAAGGA 96
Qy      1423 GGGCGCTTGCACGCTGAGATTCATGACTGGGCTG 1457
Db      95  GGGGGGTGCGAAGCGAGGCTAGTGACTGAGATG 61

RESULT 17
BZ685785 899 bp DNA linear GSS 05-FEB-2003
LOCUS PUBT547D ZM 0.6 1.0 KB Zea mays genomic clone ZMBRta032112,
DEFINITION genomic survey sequence.
ACCESSION BZ685785
VERSION BZ685785.1 GI:28245311
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACMD
clade; Panicoideae; Andropogonae; Zea.
1 (bases 1 to 899)
Whiteley,C.A., Quackenbush,J., Van Aken,S., Utecherback,T.,
Bennett,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
COMMENT TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteley@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
Location/Qualifiers
source 1..899
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBRta032112"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN
Query Match 40.6%; Score 592; DB 9; Length 899;
Best Local Similarity 80.3%; Pred. No. 2.7e-171;
Matches 742; Conservative 0; Mismatches 155; Indels 27; Gaps 3;

Qy      318  AGCAGCAGTGGGGAATTTTGAACAATGGCGCAGAGCTATCCAGCAATGCCGCTGAG 377
Db      899  AGCAGCAGTGGGGAATTTTGAACAATGGCGCAGAGCTATCCAGCAATGCCGCTGAG 840
Qy      378  TGAAGAAAGCTTGGGTTGTAAGCTCTTTCAGTGCAGAAAGAAAGTTACGTAATA 437
Db      839  TGAAGAAAGCTTGGGTTGTAAGCTCTTTCAGTGCAGAAAGAAAGTTACGTAATA 797
Qy      438  ATCGTGAATGACGCTATGACGTAAGAAAGCAGCGCTAACGAGTGCAGAGCGCG 497
Db      796  -----TATATGCGGTACCGGAAGATTAAGCCCGGCTTACTTCTGTCAGAGCGCG 745
Qy      498  GGTATATGTAAGGTGCGAAGCGTTAATCGAATTAATTAATGCGTAAAGGCTGCGAGCGG 557

```

```

Db      744  GGTATATGAAAGGGGGCTAGCGTTGCTCGAATCACTGGCGTAAAGGGCGGTAGCGG 685
Qy      558  CTTGTAAATGATGTGAATTCGCCGGGCTTAACTTGGGAATTCGCTTGAATACTACA 617
Db      684  CCAATCAAGTCGGGGGTGTAAGCCGTGCTCAACCAAGAAATGCTTTCATCTGTTT 625
Qy      618  GCGTGAAGTGGGCGAGAGGAGTGAATTCATGTGTAGAGAGTAAATGCGTAGAGATA 677
Db      624  GCGTGAAGTGGGCGAGAGTGAATTCATGTGTAGAGAGTAAATGCGTAGAGATA 565
Qy      678  TGAAGAAATCATGATGAGCGAAGCGACCTCTGGGTTAACTGACGCTCATGACGAA 737
Db      564  CCGAAGAACACAGTGGGGAAGGGCGGCACTGACCAACTGACGCTGAGCGCGAA 505
Qy      738  GCGTGGGAGCAACAGATTAATGATCCTGTGTATCCAGCGCCCTTAACGATGTAAC 797
Db      504  GCGTGGGAGCAACAGATTAATGATCCTGTGTATCCAGCGCCCTTAACGATGTAAC 445
Qy      798  GTTGTGGGCTTATTAAGCTTGTGAAGAGCTAACGCTGAAGTTGACCGCTGGGGA 857
Db      444  GCTGTGGGCTTATTAAGCTTGTGAAGAGCTAACGCTGAAGTTGACCGCTGGGGA 385
Qy      858  GTACGCTGCAAGATTTAAACTCAAGGAATTAAGCGGGAACCGCAAGCGGTGATTA 917
Db      384  GTACGCTGCAAGATTTAAACTCAAGGAATTAAGCGGGAACCGCAAGCGGTGATTA 325
Qy      918  TGTGATTAATTCATGATGCAACGCGAATACTTACCTTACCTTGTACATGTAGCAAT 977
Db      324  TGTGATTAATTCATGATGCAACGCGAATACTTACCTTACCTTGTACATGTAGCAAT 265
Qy      978  CTAGAGAT-AGATTAGTGTGGGGAAGCGTAAACAGGTGCTGACGTGCTGCGAC 1036
Db      264  GGAAGATCCGGGCTTCCCTTGGGGAACGCAACAGGTGCTGACGTGCTGCGAC 205
Qy      1037  TCGTGTGATGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGTATTATTTGCC 1096
Db      204  TCGTGTGATGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGTATTATTTGCC 145
Qy      1097  ATCATTTGTTGGGCACTTTAATGAGACTGCGGTGCAAAACCG-GAGAAAGTGGGAT 1155
Db      144  ATCATTTGTTGGGCACTTTAATGAGACTGCGGTGCAAAACCGGAGAAAGTGGGAT 85
Qy      1156  GACGTAAATCTCTCATGCGCCCTTATGGGTAGGGCTTCAACGTAATACAAAGCGGTAC 1215
Db      84  GACGTAAATCTCTCATGCGCCCTTATGGGTAGGGCTTCAACGTAATACAAAGCGGTAC 25
Qy      1216  AGAGGGTTGCCAACCGCGAGGGG 1239
Db      24  AGTGGACGCGAATCCCGAGGTG 1

RESULT 18
BZ440868 974 bp DNA linear GSS 13-DEC-2002
LOCUS BONKTS1TR_BO_1.6_2_KB tot Brassica oleracea genomic clone BONKTS1L,
DEFINITION genomic survey sequence.
ACCESSION BZ440868
VERSION BZ440868.1 GI:26698161
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 974)
Aylee,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utecherback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
PUBMED 15805490
COMMENT Other_GSSs: BONKTS1TF
Contact: Chris Town

```



TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtowm@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

FEATURES  
Source  
Location/Qualifiers

1..974  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONKTS1"  
/clone\_1ib="BO\_1.6.2\_KB\_tot"  
/note="Vector: PHOS1; Site 1: BclXI; 1.6-2 kb sheared  
total DNA inserted into PHOS1 using BclXI linkers"

## ORIGIN

Query Match 40.3%; Score 587; DB 9; Length 974;  
Beet Local Similarity 77.5%; Pred. No. 9.7e-170;  
Matches 760; Conservative 0; Mismatches 195; Indels 26; Gaps 3;  
237 CTACCAAGCGCAGTACGTAAGTGTGAGAGAGACGACCACTGGAGCTGAGA 296  
17 CTTACAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 76  
237 CACGCGCCAGATCTCTACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 356  
77 CACGCGCCAGATCTCTACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 136  
357 GATCCAGCAATGCCCGGTGATGAGAAAGGCTTGGGTTGAAAGCTTTCACTGAG 416  
137 GACCGAGCAATGCCCGGTGATGAGAAAGGCTTGGGTTGAAAGCTTTCACTGAG 196  
417 AAGAAAGGTTACGTTAATATCTGATCTGACGAGTATGACAGAAAGAACCGGCG 476  
197 AAGA-----AGCAATGACGGTATCTGGGAAATTAAGATCGGC 233  
477 TAACTACGTGCGAGCGCGCGGTATATCGTGGGTGCAAGGCTTAATCGAATTAATCG 536  
234 TAACTACGTGCGAGCGCGCGGTATATCGAAGGATGCAAGGCTTAATCGAATTAATCG 293  
537 GCGTAAAGGTTACGTTAATATCTGATCTGACGAGTATGACAGAAAGAACCGGCG 596  
294 GCGTAAAGGTTACGTTAATATCTGATCTGACGAGTATGACAGAAAGAACCGGCG 353  
597 GAATTGCGTTGAACTCAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 656  
354 GACAGGCGGTGAACTCAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 413  
657 GCAGTGAATGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 716  
414 GCGGTGAATGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 473  
717 CACTGACGCTCAATGCAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 776  
474 CACTGACGCTCAATGCAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 533  
777 GCGCTTAAGATGCAATGCTGTTGGGCTTATTAGGCTTG--GTAAAGCACTTAAC 834  
534 AGCGTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 593  
835 GCGTGAATGCAATGCTGTTGGGCTTATTAGGCTTG--GTAAAGCACTTAAC 894  
594 GCGTGAATGCAATGCTGTTGGGCTTATTAGGCTTG--GTAAAGCACTTAAC 653  
895 GGAACCGGAGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
Db 894 TAAAGCGGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 953  
Qy 1134 CAAACCGGAGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1193  
Db 894 TAAAGCGGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 953  
Qy 1194 CACGTAATGCAATGAGCGGCTA 1214  
Db 954 CACGTAATGCAATGAGCGGCTA 974

## RESULT 19

DN469396/C DN469396 756 bp mRNA linear EST 01-AUG-2005

LOCUS USDA-FP\_143378 Diaphorina citri Kuwayama (Hemiptera: Psyllidae) 1073

DEFINITION Diaphorina citri cDNA clone WHDc048\_C03 5', mRNA sequence. 1073

ACCESSION DN469396.1 GI:71538653 1073

VERSION DN469396.1 GI:71538653 1073

KEYWORDS EST. 1073

SOURCE Diaphorina citri 1073

ORGANISM Diaphorina citri 1073

REFERENCE Hunter, W.B., Dang, P.M. and McKenzie, C.L. 1073

AUTHORS Hunter, W.B., Dang, P.M. and McKenzie, C.L. 1073

TITLE Gene expression in adult Asian citrus psylla, Diaphorina citri 1073

COMMENT Unpublished (2006) 1073

JOURLN Contact: Wayne Hunter, Phat Dang, USDA, ARS 1073

USDA-ARS U.S. Horticultural Research Lab, Subtropical Insect Research Unit 1073

2001 South Rock Road, Ft. Pierce, FL 34945, USA 1073

Tel: (772) 462-5898, (772) 462-5940 1073

Fax: (772) 462-5986 1073

Email: whunter@ushrl.ars.usda.gov 1073

Seq primer: T3 Primer. 1073

Location/Qualifiers 1073

1..756 1073

/organism="Diaphorina citri" 1073

/mol\_type="mRNA" 1073

/db\_xref="taxon:121845" 1073

/clone="WHDc048\_C03" 1073

/sex="Mixed genders" 1073

/tissue\_type="whole body" 1073

/dev\_stage="adults" 1073

/lab\_host="X11-Blue" 1073

/clone\_1ib="Diaphorina citri Kuwayama (Hemiptera: 1073

Psyllidae)" 1073

/note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: 1073

/note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: 1073

/note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: 1073

/note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: 1073

/note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: 1073

/note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: 1073

/note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: 1073

/note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2: 1073





Db 497 GTCCAGCAGCCCGCGTAAATACGTAGGGTGCAGAGCGTTAATCGGAATTACTGCGGCTAAA 556  
Qy 544 GGGTGCAGAGCGCGCTTTGTAACTGAGATGTGAAATCCCGGGCTTTAACTTGGGAATTGC 603  
Db 557 GCGTCCGAGCGCGTTATGTAAAGACAGATGTGAAATCCCGGGCTTCAACTTGGGACCTGCG 616  
Qy 604 GTTTGAAACTTCAAGAGCTTAGAGTGTGCGAGAGGGAGTGGAAATTCATGTGTAGAGTGA 663  
Db 617 ATTGTGACTGATAGCTAGAGTACGTAGAGGGGAGTGAATTCGGGTGTAGAGTGA 676  
Qy 664 AATGCGTAGATATAGAAACATCGATGCGGAAGAGCGCTCTGAGTTAACTGAC 723  
Db 677 AATGCGTAGATATGCGAGAAACACGATGCGGAAGCAATCCCTGAGCGTGTATGAC 736  
Qy 724 GCTCATGCAAGAAAGCTGGGGAGCAAAACAGATTAGTACCTGTGTGTC 774  
Db 737 GCTCATGCAAGAAAGCTGGGGAGCAAAACAGATTAGTACCTGTGTGTC 787

RESULT 21  
BH651765 951 bp DNA linear GSS 19-FEB-2002  
LOCUS BOMFE30TR BO 2.3 KB Brassica oleracea genomic clone BOMFE30,  
DEFINITION genomic survey sequence.  
ACCESSION BH651765  
VERSION BH651765.1 GI:18709758  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica  
1 (bases 1 to 951)  
Ayele, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Utechtack, T. R., Wortman, J. R., White, O. R. and Town, C. D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
JOURNAL PUBMED 15805450  
COMMENT Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

FEATURES  
Source location/Qualifiers  
1..951  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO100DH3"  
/db\_xref="taxon:3712"  
/clone="BOMFE30"  
/clone\_1lb="BO\_2.3 kb"  
/note="Vector: pHOST1; Site\_1: BstXI; 2-3 kb sheared  
genomic DNA inserted into pHOST1 using BstXI linkers"

ORIGIN  
Query Match 39.4%; Score 574; DB 9; Length 951;  
Best Local Similarity 77.8%; Pred. No. 1e-165;  
Matches 731; Conservative 0; Mismatches 205; Indels 4; Gaps 3;

502 ATACCTAGGGTGCAGACCGTTAATCGGAATTACTGAGCGCTTAAAGGATGCGCAGCGCGCTTT 561  
Db 6 ATACAGAGAGATCCAGCGTTATCCGGAATGATTTGGGCGTAAAGCGTCTAGTGGCTTT 65  
Qy 552 GTAAGCTAGAGTGAATATCCCGGGCTTAACTGGGAATTCGTTTGAACCTACAGAGCT 621  
Db 66 TTAAGTCGCGCCGTCAATCCAGAGGCTCAACCTTGACAGCGCGGTGAAGAACTTCAAGCT 125  
Qy 622 AGAGTGTGCGAGAGGAGGTGAATTCATGTGTAGCACTGAAGTGAATGCGTAGATATGGA 681

Db 126 TGAATTAGGTAGGGGCGAGAGGAAATTCGCGTGAAGGGGTGAATGCGTAGAGATCGGAA 185  
Qy 682 AGAATCATGATGCGGAAGGACGCTCTGAGTTAACTGACGCTTACAGCAAGAGCT 741  
Db 186 AGAATCAGCAAGCGGGAAGGACCTCTGCTGGCCGACACTGACATGAGAGCAAGAGCTA 245  
Qy 742 GGGAGCAAAACAGATTAGATATACCTGTGTAGTCCAGCGCCCTTAAAGATGTCACTAGTTG 801  
Db 246 GGGAGGCAATGGGATTTAGATATCCCGAGTAGTCTTAGCCGTAAACGATGATATCTAGCG 305  
Qy 802 TTGGGCTTTATTAGCTTG--GTAAAGAGCTAAACCGGTGAAGTGAACCGCTGGAGT 859  
Db 306 CTGTGCGATGACCCCGTGAGCTGTGTAGCTAACGCTTAAATGATCCCGCTGGGAGT 365  
Qy 860 ACGTCCGAGATTAAACTCAAGAGAAATTGACCGGAGACCCGCAAGAGCGGTGATTAATG 919  
Db 366 ACGTTCGCAAGAAAGAACTCAAGAGAAATTGACCGGAGCCCGCAAGAGCGGTGAGCATG 425  
Qy 920 TGAATTATTTGATGCAACGCGAAACCTTACCTACCTTGAATGAGGAATTTTCT 979  
Db 426 TGGTTTATTTGATGCAAGCGGAAGAACTTACCAAGGCTTGAATGCGCGGAATCTCT 485  
Qy 980 AGAGATGATTAATGTC-CTTCGGAACCGCTAACAGAGGTGTCATGCTGTGTCAGTCT 1038  
Db 486 TGAAGAGAGGGGTGCTTCGGGAACCGGACACAGGTGTGATGCTGTGTCAGTCT 545  
Qy 1039 GTGTGTCGATGATGTTGGGTTAAGTCCCGCAACAGAGCCCAACCTTGTATTAATTCGCAT 1098  
Db 546 GTGCGGTAAAGGTGTGGGTAAAGTCCCGCAACAGAGCCCAACCTTGTGTAAATGTTGCAC 605  
Qy 1099 CATTGGTGTGGCACTTTAATGAGACTGCGGTGACAAACCGGAGAAAGTGGAGATGAC 1158  
Db 606 CTTGATGTTTGAACCTGAAACAGATCTCCGCTGTATAGCCGAGAAAGTGAAGATGAC 665  
Qy 1159 GTCAAGTCCATGATGCGCTTATGAGGTAGGCTTCAACAGTAATCAATGCGCTTACAGA 1218  
Db 666 GTCAAGTATCATGCGCTTATGCTGCGGAGACACAGTCTCAATATGCGCGGAGACAA 725  
Qy 1219 GGGTTCGCAACCCCGAGAGGGAGCTAATCTCAAGAAAGCGGTGTGTCGCGATCGAG 1278  
Db 726 GGGTTCGCAATCCCGAGAGGGTGAAGCTTAACTCAAGAAACCGGTCTTCAAGTGTGATGAG 785  
Qy 1279 TCTGCAACTGCACTCCGGAAGTGCAGATGCTGTAATGCGGATGAC-CATGTGCGCG 1337  
Db 786 CTGCAACTGCTCCCTGTCATGAAGCCGGAATGCTGTATGCTCCGCTACGCATACGGCG 845  
Qy 1338 TGAATACGTTCCCGGCTTGTAGACACCGCCGCTCACACCATGAGAGTGGTTTACCA 1397  
Db 846 TGAATTCGTTCCCGGCTTGTAGACACCGCCGCTCACACTATGGAGCTGGCATATGCC 905  
Qy 1398 GAAAGAGTATGTAACCTTAACCGCAAGAGAGGGGTGCGCAAGG 945  
Db 906 GAAAGTATGTAACCTTAACCGCAAGAGAGGGGTGCGCAAGG 945

RESULT 22  
CK290947 851 bp mRNA linear EST 02-AUG-2004  
LOCUS EST753661 Nicotiana benthamiana mixed tissue cDNA library,  
DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NBMBY50 5',  
end, mRNA sequence.  
ACCESSION CK290947  
VERSION CK290947.1 GI:39870911  
KEYWORDS EST.  
SOURCE Nicotiana benthamiana  
ORGANISM Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
Asterales; Lamiales; Solanales; Solanaceae; Nicotiana.  
REFERENCE 1 (bases 1 to 851)  
Buell, C. R., Hart, A., Zismann, V., Karamycheva, S. A., Day, B.,  
Staskiewicz, B., Jin, H. and Baker, B.

TITLE  
JOURNAL  
COMMENT

Generation of EST sequences from *Nicotiana benthamiana*  
Unpublished (2003)  
Contact: Robin Buehl  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@igf.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: AYT TAG ACA CTA TAG.

## FEATURES

source  
1..851  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NBMBY50"  
/issue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-TonA"  
/clone\_11b="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"  
/note="Vector: PCWVSport6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from *Nicotiana benthamiana*  
tissues that include callus, roots from liquid culture  
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen  
challenged leaves (*Pseudomonas syringae* pv *tomato* 12 hr;  
*Xanthomonas campestris* pv *campestris* 12 hr, 18hr;  
*Xanthomonas syringae* pv *phaseolicola* 18hr, and *Xanthomonas*  
*campestris* pv *vesicatoria* 18hr). RNA was isolated from  
these tissues and pooled in approximately equal molar  
amounts."

## ORIGIN

Query Match 39.3%; Score 572.6; DB 7; Length 851;  
Best Local Similarity 81.0%; Pred. No. 2.7e-165;  
Matches 691; Conservative 0; Mismatches 159; Indels 3; Gaps 2;  
4 GAACGCTGCGGCGCATCTTACACATGACGAGCGAGACCGATGCTTGCATCTG 63  
1 GAACGCTGCGGCGCATCTTACACATGAGTGAAGCGAGCGGATCTTGTACTCTG 60  
64 GTGGCGAGTGGCGGCGGCTGAGTAATGCAATCGAAAGTATCCAGAAAGCGGGGTTAAAG 123  
61 GTGGCGAGCGGCGGCGGCTGAGTAATGCTTGAAGTCTGCTGATGCGGGGATTAAG 120  
124 CATCGAAAGATGCTATATCCCATATCTCTTAAGAGAGAAAGCGGGGATCGAAAGAC 183  
121 CTCGAAAGCGAGCGCTATATACCGCATATGCTCTTACGAGAGAAAGCGGGGATCTTGGGCG 180  
184 CTGGCGCTTTGGAGCGCGCATGCTGATTAAGCTGATGGGTTAAAGGCTTACCA 243  
181 CTGGCGCTTACATGAGCTGAGCTGAGTAACTAGCTAGTGAAGTAAAGGCTTACCA 240  
244 GGGCGAGATCAATGATTTGGTCTGAGAGAGACGACCACTGGGACTGAGACGCGCC 303  
241 GGGCGAGATCCGTAATCGTCTGAGAGATGATCACTGAGAACTGAGACGAGGCTC 300  
304 CAGACTCTTACGAGAGCGAGCAATGGGGAATTTTGGACATATGGGCGGACCTGATCCAG 363  
301 CAGACTCTTACGAGAGCGAGCAATGGGGAATTTTGGACATATGGGCGGACCTGATCCAG 360  
364 CAATGCGCGGTGAGTGAAGAGGCTTGGGTTGAAGCTTTTCACTGAGAGAGAAA 423  
361 CCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGAAGCTTTTCACTGAGAGAGAG 420  
424 GGTACCGTAATATATGTAATCATGACGCTATGACAGAGAAAGACCGGCTAATAC 483  
421 GAGTTACCTATATATGATGATTTTGAAGCTTACCGACAGAAATAGACCGGCTAATCTC 480  
484 GTGCGACAGACCGCGGTAATATCGAAGGTGCAAGCTTAACTCGAAGCGGCGTAAA 543  
481 GTGCGACAGACCGCGGTAATATCAGAGGGTGCAGAGCTTAAATCGAATTAATCTGGGCGTAAA 540

QY 544 GGGTGGCAGCGCGCTTTGTAGTCAGATGTGAATCCCGGCTTAACCTGGGAATTGC 603  
DB 541 --GCGCGTAGGTGTTTGTATAGTGTGATGTGAATCCCGGCTCAACCTGGGACTGC 598  
QY 604 GTTTGAACCTAACAGGCTAGAGTGTGCAAGGAGGAGTGTGAATTCATGTGTACAGTGA 663  
DB 599 ATCCAAAGCTGGCAAGCTAGAGTATGTAAGGAGGTGTGGAAATTTCCGTGTACCGGTGA 658  
QY 664 AATGCTGAGATATGTGAAGAACTCATGATGGCGAAAGCAGCTCTGGGTTAACTGAC 723  
DB 659 AATGCTGAGATATGTGAAGAACTCATGATGGCGAAAGCAGCTCTGGGTTAACTGAC 718  
QY 724 GCTATGACAGAAAGCGTGGGAGCAAAACAGATTAGTACCTGTAGTCACGCCCTA 783  
DB 719 ACTGAGTGCAGAAAGCGTGGGAGCAAAACAGATTAGTACCTGTAGTCACGCCCTA 778  
QY 784 AACGATGCAACTGATTTGGGCTTATTAAG-GTTGTGTAACGAAGCTAACCGCTGAAG 842  
DB 779 AACGATGCAACTGATTTGGGAGCTTGAAGCTTGTAGTGTGCGACGCACTTAAG 838  
QY 843 TTGACCGGCTGGG 855  
DB 839 TTGACCGGCTGGG 851

RESULT 23  
LOCUS AQ957362 746 bp DNA linear GSS 28-JAN-2000  
DEFINITION LERAP36TF LERA Arabidopsis thaliana genomic clone LERAP36, genomic  
survey sequence.

ACCESSION AQ957362  
VERSION AQ957362.1 GI:6785063  
KEYWORDS  
SOURCE Arabidopsis thaliana (chale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
Rosids; eustroids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
AUTHORS 1 (bases 1 to 746)  
Buell, C.R., Jhn, X., Pal, G., Barnstead, M., Bowman, C., Uterbach, T.,  
Feldblyum, T., Liang, F., Creasy, T. and Fraser, C.M.  
Genomic survey sequencing of *landberg erecta* ecotype of  
Arabidopsis thaliana and identification of sequence-based  
polymorphisms  
Unpublished (2000)

JOURNAL  
COMMENT  
Contact: Xiaoying Lin  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: at@igf.org  
For additional information, see <http://www.tigr.org/cdb/ac/at.html>  
Seq primer: TF  
Class: shotgun.

## FEATURES

source  
1..746  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Landberg erecta"  
/db\_xref="taxon:3702"  
/clone="LERAP36"  
/clone\_11b="LERA"  
/note="Organ: Leaf; Vector: pHS01; Total genomic DNA was  
sheared to 0.9-1 Kbp before ligation."

## ORIGIN

Query Match 38.9%; Score 566.4; DB 9; Length 746;  
Best Local Similarity 86.5%; Pred. No. 2.1e-163;  
Matches 624; Conservative 0; Mismatches 97; Indels 0; Gaps 0;  
85 AGTAATGATCGGAGCTATCCAGAGAGGGGTAAACCATCGAAGATGTGCTAATAC 144  
1 AGTAATATATCGGAGCTGCTCTTGTATATGGGGATTAATAGTCAAGAAATAGCTAATAC 60

QY		145	CGCATATATCTCTAAGBAGAAAAACGCGGGGATCGAAAAGACCTTGGCGCTTTTGGAGCGGCG	204
Db		61	CGCATACGCCCTGAGGGGAAAAGTAGGGGATCTTCGACCTTACGTTATTAAGAGGCGCG	120
QY		205	ATGTCGTATTTAGCTAGTTGTGGGGTAAAGGCTTACCAAGCGACGATCAGTAGTTGGTC	264
Db		121	ATATCTGATTAGCTAGTTGTGGGGTAAATGGCTTACCAAGGCTTGATCAGTAGTGGTC	180
QY		265	TGAGAGACGACACGACCACTGGGACCTGAGACACGCGCCAGACTCTTACGGGAGGCGC	324
Db		181	TGAGAGGACGACCAAGCCACACTGGAACTGAGCACGCGTCCAGACTCTTACGGGAGGCGC	240
QY		325	AGTGGGGAATTTTGGACAAATGGGCGCAAGCCGATCCAGCAATGCGCGGTAGTGAAGAA	384
Db		241	AGTGGGGAATTTTGGACAAATGGGCGCAAGCCGATCCAGCAATTCGCGTGAAGAA	300
QY		385	GAGCCTCGGGTGTGTAAAGCTCTTTTACGTACGAGAAAGAAAGGTTACGTTAAATTCGTGA	444
Db		301	GAGCCTCGGGTGTGTAAAGCTCTTTTACGAGAGGAAAGAACTTACTTTCTAATTAAGAGTG	360
QY		445	CTCATGACGCTATCGACAGAAAGACACCGGCTTAACTACGTCGACGACCGCGGTAA	504
Db		361	AGGCTGACGCTACCTTGTGAATGAAGACACCGGCTTAACTACGTCGACGACCGCGGTAA	420
QY		505	CGTAGGGGTGCAAGCGTTATATCGGAATTAATCTGGGCGGTAAAGGGTGGCGAGGCGGCTTTGTA	564
Db		421	CGTAGGGGTGCAAGCGTTATATCGGAATTAATCTGGGCGGTAAAGGGTGGCGAGGCGGCTTTGCA	480
QY		565	AGTCAGATGTGAATCCCGCGGGCTTAACTCTGGGAATTCGTTTGAATCAACAAGGCTAGA	624
Db		481	AGTCAGATGTGAATCCCGCGGGCTTAACTCTGGGAATTCGTTTGAATCAACAAGGCTAGA	540
QY		625	GTTGTGGCAGAGGAGGTGGAATTCATGTGTACGATGGAATTCGCTAAGATATGGAAGA	684
Db		541	ATATCTCAGAGGGGGGTGAATTCACAGCTGTAGCAGTGAATTCGCTAAGATATGGAAGA	600
QY		685	ACATGATAGGGGGAAGGACGAGCCTCCTGGGGTTAACTGAACGCTCAATGACGAAAGCGTGGG	744
Db		601	ATACCAATAGGGGAAGGACGAGCCTCCTGAGATTAATTTGACGCTCATGACGAAAGCGTGGG	660
QY		745	GAGCAACAGGATTAGATACCTGTGTAGTCCAGCGCCCTTAAACGATGTCAACTAGTTGTG	804
Db		661	GAGCAACAGGATTAGATACCTGTGTAGTCCAGCGCCCTTAAACGATGTCTACTAGTTGTG	720
QY		805	G 805	
Db		721	G 721	

RESULT 24

LOCUS CN757363

DEFINITION ID00AALICE08M1 APMS Acyrthosiphon pisum cDNA clone ID00AALICE08 5', mRNA sequence.

ACCESSION CN757363

VERSION CN757363.1

KEYWORDS GI:47531286

SOURCE EST.

ORGANISM Acyrthosiphon pisum (pea aphid)

REFERENCE Hunter, W., Martinez-Torres, D., Rahbe, Y., Sabater-Munoz, B., Stern, D., Tagu, D. and Winkler, P.

AUTHORS An expressed sequence tags database for the pea aphid Acyrthosiphon pisum

TITLE Unpublished (2004)

JOURNAL INRA Rennes

COMMENT Contact: D. Tagu

UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France

Tel: +33.2.23.48.51.65

FEATURES			
source			
1. 860			
Location/Qualifiers			
/organism="Acyrtosiphon pisum"			
/mol_type="mRNA"			
/culivar="developmentstage"			
/db_xref="taxon:7029"			
/clone="ID0AA1CE08"			
/tissue_type="whole insect"			
/dev_stage="nymphs and adults (parthenogenetic females)"			
/lab_host="Xrl-Blue"			
/clone_lib="ApMS"			
/note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI; Sample name: ID0AAA; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University; Soil conditions: Soil; Sowing date: 01/06/1999; Harvesting date: 01/06/1999; Stress date: no stress; Description: Aphids inoculated on one-week old <i>Vicia faba</i> under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction.; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 °C)"			
Query Match	38.7%	Score 564;	DB 7; Length 860;
Best Local Similarity	81.5%	Pred. No. 1.2e-162;	
Matches 701;	Conservative 0;	Mismatches 155;	Indels 4; Gaps 4;
OY	353	GCCGATCCGAGCATGCCGCGGTGAGTAAGAGGCGCTCGGATTGTAAGCTCTTTCAGT	412
DB	1	GCCGATGCAAGCTATGCCGCGGTGATGAAGAGGCTTAGGATTGTAAGTACTTTTCAGC	60
OY	413	CGAAGAGAAAGTTACGTTAAATATCGTACTCATGACCGTATTCGACAGAGAAAGCAC	472
DB	61	GGGAGAGAAATAATTAATTAATTTATTTCTGACCTTACCGGCAAGAAAGCAC	120
OY	473	CGGCTAACTACGTGCCAGACGCCGCGTAACTAGTGGGTGCAAGCGTTATTCGAATTA	532
DB	121	CGGCTAACTCGGTGCCAGACGCCGCGTAACTAGGAGGTGCAAGCGTTATTCGAATTA	180
OY	533	CTGGGCGGTAAAGGTTGCGCAGCGCGCTTTGTAATCGATGTAATCCCGGCGCTTAAC	592
DB	181	CTGGGCGGTAAAGACGCGTAGTGATTTTAAAGTCAGGTGTAATCTTAGGCTTAAC	240
OY	593	CTGGGAATTTGGTTGAAATCAACAAGGCTAGTGTGCGACAGGAGGTGAAATTCATG	652
DB	241	CTAGGAATGATTTGAAATCTGAAATCTAAGTTTCTAAGGAGGTAAATTTATGAG	300
OY	653	TGTAGCAGTGAATTCGCTGAGATATGGAACAATCGATGAGCGCAAGGCGACCTCTGAG	712
DB	301	TGTAGCGGTGAATTCGCTGAGATATCTGAGAAATACCGGTGCGAAACGCGCTCTTAA	360
OY	713	TTTAACCTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTAATACCTCTGTG	772
DB	361	CGAAATCTGACACTGAGGCGCGAAAGCGCTGGGAGCAACAGATTAATATCCCTGTG	420
OY	773	TTCAAGCGCTTAAAGATGCAACTATGTTGTTGGGCTTATTAAGCTTG-GTAAAGACT	831
DB	421	TTCAATGCCGTAAGAGATGCACTTGAAGTTGTTTCCAGAGAAATGACTCTTCCAACT	480
OY	832	AACCGTGAAGTTACCGCTGGGAGATGACGCTGCAGATTTAAATCTCAAGAAATTA	891
DB	481	AAACCATTAATGATGACCGCTGGGAGATGACGCGCAAGGCTTAAATCTCAATGAATTA	540
OY	892	CGGGAGCCCGCAAGCGGTGATTAATGTGATTAATTCATGCAACGCGAAATCTTA	951

Db 541 CCGGGGCCCCGACAGCGGTGAGCATGTGTTTAATTCGATGCAACCGGAAAACTTAA 600  
Qy 952 CCTACCCCTTGACATGTAGCGAAATTTTCTAGAGATAGATTAGTG-CTTCGGGAACGCTAAC 1010  
Db 601 CTTGCTTGTGACATCCACAGAAATCTTTAGAAATTAAGAGTCCCTTCGGGAGCTGTGAG 660  
Qy 1011 ACAGGTGCTGACATGCTGTGTCAGCTGCTGTGCTGAGATGTGGTTAACTCCGCAAC 1070  
Db 661 ACAGGTGCTGACATGCTGTGTCAGCTGCTGTGCTGAGATGTGGTTAACTCCGCAAC 720  
Qy 1071 GAGCGCAACCCCTTGATCTATTAATTCGATC-ATTGCTGGGCACTTTAATGAGCTGCCG 1129  
Db 721 GAGCGCAACCCCTTGATCTATTAATTCGATC-ATTGCTGGGCACTTTAATGAGCTGCCG 780  
Qy 1130 GTCGACAAACCGGAGAAAGGTGGGATGACGTCAGATGCTCATGG-CCCTTAATGGGTAGGG 1188  
Db 781 GTTATTAACCGGAGAAAGGTGGGAGACGCTCAAGTATCATGTGCCCCCTTAACGACAGGG 840  
Qy 1189 CTTGACACGTTAATATGATGG 1208  
Db 841 CTACACACGTTGCTACATGG 860

RESULT 25  
BZ459292 977 bp DNA linear GSS 13-DEC-2002  
LOCUS BONKQ28TF BO 1.6.2 KB tot Brassica oleracea genomic clone BONKQ28,  
DEFINITION genomic survey sequence.  
ACCESSION BZ459292  
VERSION BZ459292.1 GI:26740661  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eutroids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 977)  
Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Uteback,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
JOURNAL PUBMED 15805490  
COMMENT Other GSSs: BONKQ28TR  
Contact: Chris Town

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.  
FEATURES  
source  
1..977  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONKQ28"  
/clone\_id="BO\_1.6.2\_KB\_tot"  
/note="Vector: pHD51; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pHD51 using BstXI linkers"

## ORIGIN

Query Match 38.6%; Score 562.2; DB 9; Length 977;  
Best Local Similarity 75.7%; Pred. No. 4,6e-162;  
Matches 746; Conservative 0; Mismatches 213; Indels 26; Gaps 3;  
Qy 347 GCGCAAGCCTGATCCAGCAATGCCGCTGATGAGAAGGCTTCGGTTGTAAGCTCT 406  
Db 1 GCGAAAGCCTGATCCAGCAATGCCGCTGATGAGAAGGCTTCGGTTGTAAGCTCT 60

Qy 407 TTCACTCGAAGAAAGATTACGGTAATAATCGTACCTGACTGATCGATGACAGAA 466  
Db 61 TTCCCAAGAAAGAAAGCA-----ATGACGGTATCTGGGGAAAT 97  
Qy 467 AAGCACCAGCTTAATACATGTCAGACAGCCGCGGTAAATCGTAGGGTGCAGCGTTAATCG 526  
Db 98 AAGCATCCGCTAACTCTGTGACAGCAGCCGCGGTAAATCAGAGGATGACAGCGTTATCCG 157  
Qy 527 GAATTAAGTGGCGTAAAGGGTCCAGCGCGCTTTGTAGTCAAGATGTGAATCCCGGG 586  
Db 158 GAATTAAGTGGCGTAAAGGGTCTGTAGTGGCTTTTAAATCGCCGCTAAATCCAGGG 217  
Qy 587 CTTAATCCTGGGAATTGCCGTTTGAATCTAAGAGTGTGGGACAGAGGAGTGAAT 646  
Db 218 CTCAACCCCTGACAGCGCGGTGAATCTAAGCTTGAATACGATAGGGGACAGAGGAAT 277  
Qy 647 TCCATGTGTAGAGTGAATGCGTGAAGATATGAGAAACATCGATGGCGAAGCAGCT 706  
Db 278 TTCGCGTGAAGCGGTGAATGCGTGAAGATGGAAGAACACCAACGCGGAAGCAGCT 337  
Qy 707 CTTGGGTAAACCTGACGCTCATGACGAAACGCTGGGAGACAAACAGATTAGATACC 766  
Db 338 GCTGGGCGACACTGACACTGAGACGAAACCTAGGGAGGAAATGGATTAGATACC 397  
Qy 767 TGTAGTCCACGCCCTTAACGATGTCACTAGTTGTGGCCCTTAATAGGCTTG--GTAA 824  
Db 398 CAGTGTCTTCAACCCGTAACGATGATATAGCGCTGTGCTATTCGACCCGTCGATGC 457  
Qy 825 CGAAGCTAACGCGGTGATGATGACCCGCTGGGAGTACGCTGCAAGATTAAACCTCAAG 884  
Db 458 TGTAGTAAACGCGTTAAGTATCCCGCTGGGAGTACGTTCCAGAAATGAACCTCAAG 517  
Qy 885 GAATTAAGCGGGACCCGACAAACCGGTGATATGTGATTAATTCATGACACCGGAA 944  
Db 518 GAATTAAGCGGGACCCGACAAACCGGTGATATGTGATTAATTCATGACAAAGGAAG 577  
Qy 945 AACCTTAACCTACCTTGAATGATGAGAAATTTCTAGAGATGATTAATG-CTTCGGGA 1003  
Db 578 AACCTTAACGAGGCTTGAATGATGACCGGAAATCTCTTGAAGAGAGGGGTGCTTCGGGA 637  
Qy 1004 CGCTAACAGGCTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063  
Db 638 CCGGACACAGGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 697  
Qy 1064 CCGCAACGAGCGCAACCTTGTCTATTAATTCGATATTTGGTGGCACTTTAATGAGA 1123  
Db 698 CCGCAACGAGCGCAACCTTGTCTATTAATTCGATATTTGGTGGCACTTTAATGAGA 757  
Qy 1124 CTGCGGTGACAAACCGGAGAAAGGTGGGATGATGATGATGATGATGATGATGATGAT 1183  
Db 758 CTGCGGTGATTAAGCCGGAAGAGGTGAGATGACGTCAAGTATCATGCTCCCTTAATGCC 817  
Qy 1184 TNGGCTTCAACGATTAATCAATGAGCGGTAAGAGAGGTTGCAACCCCGGAGGAGAG 1243  
Db 818 CTGCGGACACACGCTCTAATATGCGCGGACAAAGGTCGCGATCCCGGAGGTGAGC 877  
Qy 1244 TAATCTCAGAAAGCGGTGTAGTCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 1303  
Db 878 TAATCTCAGAAAGCGGTGTAGTCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 937  
Qy 1304 GAATCGTATTAATTCGCGATGAC 1328  
Db 938 GAATCGTATTAATTCGCGATGAC 962

RESULT 26  
BZ464837 947 bp DNA linear GSS 13-DEC-2002  
LOCUS BONKP76TR BO 1.6.2 KB tot Brassica oleracea genomic clone BONKP76,  
DEFINITION genomic survey sequence.  
ACCESSION BZ464837  
VERSION BZ464837.1 GI:26752194  
KEYWORDS GSS.

Source	Organism	Reference Authors	Title	Journal	PubMed	Comment
Brassica oleracea	Brassica oleracea	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.				
1 (bases 1 to 947)	Ayle, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utecherbeck, T.R., Wortman, J.R., White, O.R. and Town, C.D.	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis				
Genome Res. 15 (4), 487-495 (2005)	15805490					
Other_GSSs: BONKP76TF	Contact: Chris Town					
TIGR	9712 Medical Center Drive, Rockville, MD 20850, USA.					
Tel: 301-838-3523						
Fax: 301-838-0208						
Email: cdtown@tigr.org						
DNA is from a doubled haploid provided by Tom Osborn.						
Seq primer: 1R						
Class: sheared ends.						
Location/Qualifiers						
1..947						
/organism="Brassica oleracea"						
/mol_type="genomic DNA"						
/strain="TOL00DH3"						
/db_xref="taxon:3712"						
/clone="BONKP76"						
/clone_1lb="BO_1.6.2_KB_tot"						
/note="vector: phos1; site_1: Bext1; 1.6-2 kb sheared total DNA inserted into phos1 using Bext1 linkers"						
ORIGIN						
Query Match	38.5%;	Score 560.4;	DB 9;	Length 947;		
Best Local Similarity	76.8%;	Pred. No. 1.7e-161;				
Matches 734;	Conservative 0;	Mismatches 196;	Indels 26;	Gaps 3		
166 AGCAGGGGATCGAAGACCTTGCGCTTTTGGAGCGGCCGATGCTGATTAGTACTTGCT						
6 AGCTGAGGACAAAGAGGATCCGCCGAGGAGGCGCTCGCTCTGATTAGTTACTTGCT						
226 GGGGTAAAGGCTTACCAAGGCGACATCATGTAAGTTGGTTGAGAGAGACACAGCCAC						
66 GAGGCAATAGCTTACCAAGCGCATGATCATGTGCTGCTCCAGAGGATGATCAGCCAC						
286 TGGGACTGAGACAGGCCACAGCTCCTACGGAGGACAGAGGGGGAATTTTGGACAATG						
126 TGGGACTGAGACAGGCCACAGCTCCTACGGAGGACAGAGGGGGAATTTTCCGCAATG						
346 GCGCGAACCTGATCAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGCTC						
186 GCGGAAACCTGAGCGAGGAATGCCGCGTGAAGTGAAGAGGCTTACCGGTTCCGAATTC						
406 TTTCAGTGAAGAAAGGTTTACGGTAATATGTTGTAATGATGATGACGTTATGACAGAA						
246 TTTTCCAGAGAAAGG-----AGCAATGACGGTATCTGGGGA						
466 GAAGACAGGGCTAACTAGCTGCACAGCGCGGTAATCTGTAAGGTTGCAAGCGTTAATC						
283 TAGCATCGGCTTACTCTGTGCGACAGCGCGGCTAATCAAGATGCAAGCGTTATCC						
526 GGAATTACTGGGCGTAAAGGATGCGCAGCGCGCTTTGTAAGTCAAGTGTGAATCCCGG						
343 GGAATGATTTGGGCGTAAAGCGTCTGTAAGTGGCTTTTAAAGTCCGCGTCAATCCAGG						
586 GCTTAACCTGGGAATTGGCTTTGAATCAAGGCTAGAGTGTGCGACAGGAGGTGGA						
403 GCTCAACCTGGACAGCGGTGGAATCACTCAAGCTTGAAGTCCGTAAGGGGACAGAGGA						
646 TTCCATGCTGTAAGTGAATGCGGTGAAGATATGGAAGAACTTCATGGCGAAAGGACGC						
463 TTTCCGATGAGCGGTGAATGCGGTGAAGATGGAAGAACTCAAGCGCGTAAGGACCTC						

Query Match	Score	DB	Length
38.4%	559.4	DB 9	901

Best Local Similarity 78.6%; Pred. No. 3.3e-161;  
Matches 694; Conservative 0; Mismatches 186; Indels 3; Gaps 2;

```

QY 448 ATGACGGTATCGACGAAGAAGCACCAGTAACTACGTCCAGCAGCGCGGGTAAATACGT 507
DB 883 ATGACGGTATCGGGGAATACATCGCTAACTCTGTCCAGCACGCCGGTAAATACAG 824
QY 508 AGGATGCAAGCGTTAATCGGAATTAATCTGGGCGTAAAGGGTGGCAAGCGGCTTTTGAAT 567
DB 823 AGGATGCAAGCGTTAATCGGAATTAATCTGGGCGTAAAGGGTGGCAAGCGGCTTTTGAAT 764
QY 568 CAGATGTGAATATCCCGGGCTTAACCTGGGAATTCGCTTTGAATTAATCAAGCTAGAGTG 627
DB 763 CCGCGCTCAATATCCAGGGCTCAACCTGACAGCGGGTGAAGAACTACCAAGCTTGAAGTA 704
QY 628 TGGCAGAGGAGAGTGAATTTCCATGTGTGACGTGAATATCGTAGAGATATGAAGAACA 687
DB 703 CGGTAGGGGCAAGGGAAATTTCCGTGGAGCGGTGAATATCGTAGAGATATGAAGAACA 644
QY 688 TCGATGGCGAAGGACAGCTCTGGGTTAACTGACGTCAATGATGCAAGAAAGCGTGGGAG 747
DB 643 CCAACGGCGAAGAGCACTCTGCTGGGCGGACACTGACATGAGAGAGAAAGCTAGGGAG 584
QY 748 CAAACGAGATTAGATACCTGTGATGCAAGCGCTTAACGATGTCAATGTTGTTGGC 807
DB 583 CGAATGGGATTAGATACCTGTGATGCAAGCGCTTAACGATGTGATGATGATGATGATGATG 524
QY 808 CTTATTAAGCTTG--GTAAAGAGCTTAAGCGGTGAAGTTGACCGGCTGGGGAGTACGGTC 865
DB 523 GTATCACCCTGTGCACTGTGTGATGACGCGGTGAAGTATCCGCTGGGGAGTACGGTC 464
QY 866 GCAAGATTAAATCTCAAGGAATTGACGGGAGCCGACAGCGGGTGAATTTATGTAAT 925
DB 463 GCAAGATTAAATCTCAAGGAATTGACGGGAGCCGACAGCGGTGAAGTATGTAAT 404
QY 926 AATTGATGCAAGCGCAAAACCTTAACCTTGAATGATGATGATGATGATGATGATGATGATG 985
DB 403 AATTGATGCAAGCGCAAAACCTTAACCTTGAATGATGATGATGATGATGATGATGATGATG 344
QY 986 AGATTATGTG--CTTCGGGAAGCTTAACAGAGTGTGATGATGATGATGATGATGATGATGATG 1044
DB 343 AGAGGGGTGCTTCGGGAAGCTTAACAGAGTGTGATGATGATGATGATGATGATGATGATG 284
QY 1045 TGAGATGTTGGGTTAATGATCCCGCAACGAGCGCAACCTTGTCAATTAATTTCCATCAAT 1104
DB 283 TAAAGTGTGGGTTAATGATCCCGCAACGAGCGCAACCTTGTCAATTAATTTCCATCAAT 224
QY 1105 GTTGGGCACTTAAATGATGATCCCGCAACGAGCGCAACCTTGTCAATTAATTTCCATCAAT 1164
DB 223 GTTGGGCACTTAAATGATGATCCCGCAACGAGCGCAACCTTGTCAATTAATTTCCATCAAT 164
QY 1165 TCTCATGAGCGCTTAATGATGATCCCGCAACGAGCGCAACCTTGTCAATTAATTTCCATCAAT 1224
DB 163 TCAATATGAGCGCTTAATGATCCCGCAACGAGCGCAACCTTGTCAATTAATTTCCATCAAT 104
QY 1225 CCAACCGCGAAGGAGGAGCTTAATGATGATCCCGCAACGAGCGCAACCTTGTCAATTAATTTCCATCAAT 1284
DB 103 CCAATGCGCGAAGGAGGAGCTTAATGATGATCCCGCAACGAGCGCAACCTTGTCAATTAATTTCCATCAAT 44
QY 1285 ACTGCACTCCGTAAGTGCAGATGCTAGTAAATGCGCGATCAG 1327
DB 43 ACTGCACTCCGTAAGTGCAGATGCTAGTAAATGCGCGATCAG 1

```

RESULT 28  
BZ469058 1106 bp DNA linear GSS 13-DEC-2002  
LOCUS BZ469058 BO.1.6.2\_KB\_tot Brassica oleracea genomic clone BONHE54,  
DEFINITION genomic survey sequence.  
ACCESSION BZ469058  
VERSION BZ469058.1 GI:26764665  
KEYWORDS GSS.  
SOURCE Brassica oleracea

ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosid II; Brassicales; Brassicaceae; Brassica.  
REFERENCE  
AUTHORS Ayele, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Ullrich, T. R., Mortimer, J. R., White, O. R., and Town, C. D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
JOURNAL PUBMED 15805490  
COMMENT Other GSSs: BONHE54TR  
Contact: Chris Town  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: ctdow@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: 1P  
Classes: sheared ends.  
FEATURES  
source location/Qualifiers  
1..1106  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="T0100DH3"  
/db\_xref="taxon:3712"  
/clone="BONHE54"  
/clone\_1b="BO.1.6.2\_KB\_tot"  
/note="Vector: pHD51; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pHD51 using BstXI linkers"

ORIGIN  
Query Match 38.2%; Score 556.4; DB 9; Length 1106;  
Best Local Similarity 74.8%; Pred. No. 3e-160;  
Matches 812; Conservative 0; Mismatches 221; Indels 53; Gaps 7;

```

QY 3 TGAACGCTGGCGCATGCTTTAACAATGCAATGCAAGCGGACAGCGATGCTTGATCT 62
DB 1057 TGAACGCTGGCGCATGCTTTAACAATGCAATGCAAGCGGACAGCGATGCTTGATCT 1010
QY 63 GGTGGCAATGGCGGACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 122
DB 1009 -GTTTCAGTGGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 952
QY 123 GCATCGAAGATGCTTAATTAACCGCATATCTTAAGAGAAAGCAGGCGATCGAAGA 182
DB 951 AGCTGAAACGCTGCTAATTAACCGCATATCTTAAGAGAAAGCAGGCGATCGAAGA 900
QY 183 CTTGGGCTTTTGGAGCGGCGCATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 242
DB 899 -CGCCGAGGAGGAGGCGCTGCGTGTATTAATTAATTAATTAATTAATTAATTAATTAAT 845
QY 243 AGCGCAAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 302
DB 844 AGCGCAAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 785
QY 303 CCAAGCTCCCTAAGGAGGAGCGAGCATGATGATGATGATGATGATGATGATGATGATGATG 362
DB 784 CCAAGCTCCCTAAGGAGGAGCGAGCATGATGATGATGATGATGATGATGATGATGATGATG 725
QY 363 GCAATGCGCGTGAAGAGAGGCTTTCGGGTTGAAGCTTTTCACTGAGAGAGAA 422
DB 724 GCAATGCGCGTGAAGAGAGGCTTTCGGGTTGAAGCTTTTCACTGAGAGAGAA 667
QY 423 AGGTTACGGTAATTAATCGATCACTGACGTGATGACGAGAGAGAGAGAGAGAGAGAGAG 482
DB 666 -AGCAATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
QY 483 CGTCCAGAGAGCGCGGTAATAGTATGAGGTCGAAAGCTTAATCGAATTAATCGAGGCTAA 542
DB 627 TGTCCAGAGAGCGCGGTAATAGTATGAGGTCGAAAGCTTAATCGAATTAATCGAGGCTAA 568

```



Oy		543	AGGGGCGCAGCCGCCCTTTGTAAATGCACATGTGAATAATCCCGGGCTTAACTGAGAATTG	602
Dd		567	AAGCCTTAGTGTCCTTTTTTAAGTCGCCGCGCAATATCCAGGGCTCAAACCTTGACAAG	508
Oy		603	C GTTTGAATACTAACAAGGCTAGAGTGTGCACAGAGGATGGAATTCATGTGTAGCACTG	662
Dd		507	CGGTGGAAACTAACAAAGCTTAGTAGTACGGTAAGGGGACAGAGGAATTTCCGGTGGAGCGGTTG	448
Oy		663	AAATCGTAGAGATATGGAAGAACCATCGATGGCGGAAGCGACCTCTCTGGGTTAAACATGA	722
Dd		447	AAATGCTGAGAGATCGAAAAAGAACCAACACGCGGAACCATCTGTGGGCCGACATGA	388
Oy		723	CGCTATATCACGAAGCGGTGGGGACCAACAAGSATTAGATACCTCTGTAGTCCAAGCCCT	782
Dd		387	CAC TAGAGAGACGAAGAAGCTTAGGGGAGCCGATGGGATTAAGATACCTCAGTAGTCCTAGCCGT	328
Oy		783	AAACGATGTCAACTAGATTGTTGGGCTTATTTAGGCTTG--GTAAAGAGCTAACGCGTGA	840
Dd		327	AAACGATGAGATACTAAGGCGCTGTGCGTATCCAGCCGTGACGTGCTGTAGCTTAAGCGCTTA	268
Oy		841	AGTTGACCCGCTGGGGAGTACCGGTGCAAGATTTAAACTCAAAGSAATTGACGGGGACCC	900
Dd		267	AGTATCCCCGCTGGGAGTAGCTTGCGAAGATGAATCTAAAGSATTTGACGGGGGCC	208
Oy		901	GCA CAA GCGGTGATATATGTGATTAATTCATGCAAGCCGAAAAA CCTTACTRACCTT	960
Dd		207	GCA CA AGCGGTGAGACATGTGTTTAATTCATGCAAGCGMAAGCAACTT TACCAAGGCTT	148
Oy		961	GACATGTAGCGCAATTTTCTAGAGATFAGATTAGTG-CTTCGGGAACGCTAACACAGTGTCT	1019
Dd		147	GACATGCCCGCANATCTCTTGAAAAGAGGGGTGCTTTCGGGAACGCGGACACAGTGTCT	88
Oy		1020	GCATGCTGTGCTCAGCTCGTGTCTGTGAGATTGGGTTAAGTCCCGCAACGAGCGCAAC	1079
Dd		87	GCATGCTGTGCTCAGCTCGTGTGCGAAGAGGTGTTTAAAGTCCCGCAACGAGCGCAAC	28
Oy		1080	CCTTGT	1085
Dd		27	CCTCGT	22

  

RESULT_29	BZ474941/c	1069 bp	DNA linear	GSS 13-DEC-2002
LOCUS	BONKR5STR BO.1.6.2 KB tot Brassica oleracea genomic clone BONKR59,			
DEFINITION	genomic survey sequence.			
VERSION	BZ474941			
KEYWORDS	BZ474941.1 GI:26776372			
SOURCE	GSS.			
ORGANISM	Brassica oleracea			
	Brassica oleracea			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons;			
	Rosids; eurosoids II; Brassicales; Brassicaceae; Brassica.			
REFERENCE	1 (bases 1 to 1069)			
AUTHORS	Aylee,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Utecherack,T.R., Wortman,J.R., White,O.R. and Town,C.D. Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res. 15 (4), 487-495 (2005)			
JOURNAL_PUBMED COMMENT	15805490 Other GSSE: BONKR59TF Contact: Chris Town TIGR			
	9712 Medical Center Drive, Rockville, MD 20850, USA. Tel.: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TR Class: sheared ends.			
FEATURES	Location/Qualifiers			
source	1..1069			

```

/organism="Hirassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BONKRS9"
/clone_1lb="BO_1.6-2_KB_toc"
/note="Vector: pGHOST1; Site_1:
BacXI, 1.6-2 kb sheared
total DNA inserted into pGHOST1 using
BacXI linkers"

```

Query Match	38.2%;	Score 556.2;	DB 9;	Length 1069;
Best Local Similarity	74.2%;	Pred. No. 3.4e-160;		
Matches 805; Conservative	0;	Mismatches 228;	Indels 52;	Gaps 6

Qy		TTAAAGCTGTGGCGGATGCTTTATCAATGCAAGTGGAAACGGCAGCAGCGATGTTGATCT	62
Dp	1036	TGAACGCTCGGGCATTGCTTACACCATCAAGTCGACGGGA-----AGTGG	990
Qy		GGTGGGAGATGGCGGACGGGTGAGTAAATGCAATCGGAACGTATCCGAAGAGGGGGGTAACT	122
Dp	989	TGTTTTCAGATGGGGGACGGGTGATGTAACGCGTAAGACCTTGCTCTTGGAGGGGAACACC	930
Qy		123 GCATCGAAAGATGTGTCTAATACCGCATATCTCTAAGAGGAAAGCAGGGGATTCGAAAGA	182
Dp	929	AGCTGAAACGGCTGTCTAATATCCCGTAGGCTG--AGAGCAAAAGGAGAAATC-----	878
Qy		183 CTTTGGCTTTTGGAGCGGCGCATGTCATGATTAGCTAGTTGGGGTTAAAGGCTTAACA	242
Dp	877	-----CGCCGAGAGAGGGGCTCGGCTGATTTAGCTAGTTGGTGAAGCAATAGCTTAACA	823
Qy		243 AGCGCAGCATCACTGATTGTTGTCTGAGAGACGACACGACACTGCGGACTGAGACACGCG	302
Dp	822	AGCGCATGATCACTGATGCTGTGTCCGAGAGAAATCACCACTGGGATCTGAAACACGCG	763
Qy		303 CCAGATCTCTCAAGGGGAGGACAGATGGGGAAATTTTGGACATAGGAGCCAGCTGTATCCA	362
Dp	762	CCAGATCTCTTACGGAGGACGACATGGGGGAATTTTCCGCAATGGGCGAAAGCTGACGGA	703
Qy		363 GCATGCGCGCTGAGTGAAGAAAGGCTTTCGGGTTGTAAAGCTCTTTCAGTCGAAAGAA	422
Dp	702	GCAATCTCCCGTGGAGGTGAAGAGGCTACGGGTCTTAACCTTTTCCAGAGAAAGA--	645
Qy		423 AGGTTACGGTTAAATATGCTGATCATAGACGGTATCGACGAAGAACACCGGCTAACTA	482
Dp	644	-----AGCAATACCGTATCTGGGGATTAACATTCGGCTAATCTC	606
Qy		483 CGTGCCAGACGCGCGGTATATCGTAGGGTGAAGCGTTATCGGAATTACTGGCGTAA	542
Dp	605	TGTGCCAGACGCGCGGTATATCAGAGGATCGAAGCGTTATCCGGAATGATGGGCGTAA	546
Qy		543 AGGATCGCAGCGCGCTTTGTAAATGTCAGATGTGAATTCGCCGGCTTAACCTGGGAATG	602
Dp	545	AGGCTGTAGAGTGCGCTTTTAAAGTCCGCGCTCAAAATCCAGGGCTCAACCTTGACAG	486
Qy		603 CGTTTGAATCTCAAGGCTAGAGTGTGGCAAGGAGAGGTGAATTCATGTGAGAGTG	662
Dp	485	CGGTGAATCTCAAGCTTAGTAGTCGTATAGGGCAGAGGGAATTTCCGTTGAGAGCGGTG	426
Qy		663 AAATGCGTAGATATGGAAGAACATGATGGCGAAGGACGCTCTGGGTTAACTGA	722
Dp	425	AAATGCGTAGATATGGAAGAACACCAACGGCGCAAAAGCACTGCTGCGGCGCACACTGA	366
Qy		723 CGGTCAATGACGAAGCGTGGGGAGCAAAACGAAATTGATATCCCTGGTATGCCAGCGCT	782
Dp	365	CAGTGAAGACGAAGCTTAGGGGAGCGGAATGATATACCCCACTAGTCTTAAGCCTG	306
Qy		783 AAACGATGCACTAGTTGTTGGGCTTAATTAAGCTTG--GTAAAGAGCTAACCGGTGA	840
Dp	305	AAACGATGATATCTAGGGGCTGTGTGATTCGACCCGTCAGATGCTGTAGCTAACCGTGA	246
Qy		841 AGTTGACCGCTCGGGGAGTACGGTTCGCAAGTTAAATCTAAAGAAATTGAACGGGACCC	900
Dp	245	AGTATCCCGCTCGGGGAGTACGTTTCGCAAGAAATAAATCAAGAGAAATTGACGGGAGCCC	186



QY 901 GCACAGCGGTGATGATGATGATTAATTCAGATGACGCGAAGAAACCTTACCTTACCTT 960  
DB 185 GCACAGCGGTGATGATGATGATTAATTCAGATGACGCGAAGAAACCTTACCTTACCTT 126  
QY 961 GACATGATGACGAAATTTTCTAGATGATGATGATG-CTTGGGAGAGGCTTACACAGGTGCT 1019  
DB 125 GACATGCGCGGAATCTCTTGAAGAGAGGGGTGCTTCGGAGACGCGGACACAGGTGCT 66  
QY 1020 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079  
DB 65 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6  
QY 1080 CCTTG 1084  
DB 5 CCTCG 1

RESULT 30  
AQ957363/c 716 bp DNA linear GSS 28-JAN-2000  
LOCUS LERAP36R LERA Arabidopsis thaliana genomic clone LERAP36, genomic  
DEFINITION LERAP36R LERA Arabidopsis thaliana genomic clone LERAP36, genomic  
survey sequence.  
ACCESSION AQ957363  
VERSION AQ957363.1 GI:6785064  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 716)  
Buell, C.R., Lin, X., Pal, G., Barnstead, M., Bowman, C., Uterbach, T.,  
Feldblum, T., Liang, P., Creasy, T. and Fraser, C.M.  
Genomic survey sequencing of Landsberg erecta ecotype of  
Arabidopsis thaliana and identification of sequence-based  
polymorphisms  
Unpublished (2000)  
Contact: Xiaoying Lin  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: atc@igr.org  
For additional information, see <http://www.tigr.org/cdb/at.html>  
Seq primer: TR  
Class: shotgun.

FEATURES  
source  
Location/Qualifiers  
1..716  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Landsberg erecta"  
/db\_xref="taxon:3702"  
/clone\_id="LERAP36"  
/note="Organ: Leaf; Vector: pHOS1; Total genomic DNA was  
sheared to 0.9-1 Kbp before ligation."

ORIGIN  
Query Match 38.1%; Score 554.4; DB 9; Length 716;  
Best Local Similarity 87.0%; Pred. No. 1,1e-159;  
Matches 621; Conservative 0; Mismatches 91; Indels 2; Gaps 1;

QY 274 GACCAAGCAGCACTGGAGCTGAGACACGCCGAGCTCTACGGAGGACAGAGGGGAA 333  
DB 716 GACCAAGCAGCACTGGAGCTGAGACACGCCGAGCTCTACGGAGGACAGAGGGGAA 657  
QY 334 TTTTGAATGAGCGGCGCAAGCGCTGATCAGCAATGCGCGCTGAGTGAAGAGGCTTCGG 393  
DB 656 TTTTGAATGAGCGGCGCAAGCGCTGATCAGCAATGCGCGCTGAGTGAAGAGGCTTCGG 597  
QY 394 GTTGTAAAGCTCTTTCAGTGCAGAAAGAAAGGTTACGGTAAATATGCTGATCATGACG 453

DB 596 GTTGTAAAGCTCTTTCAGTGCAGAAAGAAAGAAACCTTACTTTCTAATAAGATGAGGCTGACG 537  
QY 454 GTATGACAGAAAGAACACCGGCTTACTTACCTGTCAGAGCGGGGTAATATGTAAGGCTG 513  
DB 536 GTACCTTATTAAGAGCACCGGCTTAACTTACCTGTCAGAGCGGGGTAATATGTAAGGCTG 477  
QY 514 CAAGGCTTAATCGGAATTAATGAGGCTGTAAGGAGTGCAGAGCGGCTTTGTAAGTCAGATG 573  
DB 476 CGAGGCTTAATCGGAATTAATGAGGCTGTAAGGAGTGCAGAGCGGCTTTGTAAGTCAGATG 417  
QY 574 TGAATATCCCGGCGCTTAACTTGGGAATTCGTTGAAATCTACAGGCTTAAAGTGTGGCAG 633  
DB 416 TGAATATCCCGGCGCTTAACTTGGGAATTCGTTGAAATCTGCAAGATGATATGTCAG 357  
QY 634 AGGAGGTGGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 693  
DB 356 AGGAGGTGGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 297  
QY 694 GCGAAGGAGCGCTCTGGGTTAACTGACGCTCATGACGAAACGTTGGGAGCAACA 753  
DB 236 GCGAAGGAGCGCTCTGGGTTAACTGACGCTCATGACGAAACGTTGGGAGCAACA 237  
QY 754 GGATTAGATACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 811  
DB 236 GGATTAGATACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 177  
QY 812 TTAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 871  
DB 176 AATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 117  
QY 872 TTTAAATCTAAAGAAATGACGCGGAGCGGACGCAAGCGGTGATTAATGATTAATGCG 931  
DB 116 TTTAAATCTAAAGAAATGACGCGGAGCGGACGCAAGCGGTGATTAATGATTAATGCG 57  
QY 932 ATGCAACGCGAAGAAACCTTACCTTACCTGATGATGATGATGATGATGATGATGATGAT 985  
DB 56 ATGCAACGCGAAGAAACCTTACCTTACCTGATGATGATGATGATGATGATGATGATGAT 3

RESULT 31  
B2437960 918 bp DNA linear GSS 13-DEC-2002  
LOCUS BONNV72TF\_BO.1.6\_2\_KB\_tot Brassica oleracea genomic clone BONNV72,  
DEFINITION genomic survey sequence.  
ACCESSION B2437960  
VERSION B2437960.1 GI:26692017  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 918)  
Aylee, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Uterbach, T.R., Mortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490  
Other\_GSSs: BONNV72TR  
Contact: Chris Town

JOURNAL  
PUBMED  
COMMENT

FEATURES  
source  
Location/Qualifiers  
1..918  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"



```

Db      | 284 | CCAACGGCGAAGGCACTCTGCTGGGCGGCACTGACATGAGAGAGAAAGCTTAGGGAG 343
Qy      | 748 | CAAACAGATTGATATCCCTGTAAGTCACGGCCCTAAAGATGTCAACTAGTTGTGGGCG 807
Db      | 344 | CAAATGGAGTATGATACCCAGAGTCTAGCCGTAAAGATGATATCTAGGCGCTGTGC 403
Qy      | 808 | CTATATAGGCTTG--GTAAAGAGCTAACGCGTGAAGTTGACCGCGCTGGGGAGTACGTC 865
Db      | 404 | GATACGACCCGCTGACGTGTGTAGTAAAGCGGTAAAGTATCCCGCTGGGGAGTACGTC 463
Qy      | 866 | GCAAGATTAAATCTCAAAGAAATTGACGGGAGCCGCAACAGCGGTGTATATGTCGATT 925
Db      | 464 | GCAAGATGAAATCTCAAAGAAATTGACGGGAGCCGCAACAGCGGTGTAGAGCAATGTGTT 523
Qy      | 926 | AATTGATGCAACGCGAAGAACTTACCTTACCTTGAACATGTAGCAATTTTCTAGAGAT 985
Db      | 524 | AATTGATGCAAGCGAAGAACTTACCAAGGCTTGAACATGCGCAATCTTGAAG 583
Qy      | 986 | AGATTAGTG-CTTCGGGAAGCTAACACAGTGTCTGATGCTGTGCTGTGCTGCTGTCG 1044
Db      | 584 | AAGAGGGTGCTCTTCGGGAAGCGGAGACAGTGTGTGATGCTGTGCTGCTGCTGCTG 643
Qy      | 1045 | TGAGATGTTGGGTTAAGTCCCGACAGCGCAACCTTGTCAATTAATGCTCATTTTG 1104
Db      | 644 | TAAAGTGTGGGTTAAGTCCCGACAGCGCAACCTTGTGTAGTTGCTTGCACATTA 703
Qy      | 1105 | GTTGGGCACTTAAATGAGATGCGCGTGAACAACCGGAGAAAGTGGGGATGACGTCAAG 1164
Db      | 704 | GTGTGAAACCTTGAAACAGATGCGCGGTGATTAACCGGAGAAAGTGAAGTACGTACAG 763
Qy      | 1165 | TCCTCATGCGCCTTATGGGTAGGGCTTCAACGTAATACATGAGCGGTACAGAGGGTTG 1224
Db      | 764 | TCAATATGCCCCCTTAATGCCCCCTGAGCAGACACCTGTCTACATATGGCGGAGCAAGGGTGC 823
Qy      | 1225 | CCAACCGCGAGGGAGAGTAACTCTCAAGAAAGCGCGTGTAGTCCGGATGAGTGTGCA 1284
Db      | 824 | CGATCCCGGAGGGTGAAGTAACTCCAAAGAACCGTCTCAGTTCGATTCGATTCAGAGGCTGCC 883
Qy      | 1285 | ACTGACTCCGTGAAGTGCAGATTCGCTAATGA 1316
Db      | 884 | ACTGCTGCTGATGAAGCGGAAATCGTATGA 915

RESULT 33
LOCUS      BU279543              704 bp      mRNA      linear      EST 09-APR-2002
DEFINITION BU279543 Y. Ogihara unpublished cDNA library, wh_r Triticum aestivum cDNA clone wh3h15 5', mRNA sequence.
ACCESSION  BU279543
VERSION     BU279543.1  GI:20102041
KEYWORDS    Triticum aestivum (bread wheat)
SOURCE      Triticum aestivum
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poideae; Triticeae; Triticum.
            1 (bases 1 to 704)
REFERENCE   Ogihara, Y. and Murai, K.
            Expressed genes in Triticum aestivum
            Unpublished (2002)
            Contact: Tadenu Shin-i
            Center for Genetic Resource Information
            National Institute of Genetics
            111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tehin@genes.nig.ac.jp.
FEATURES
    source
        1..704
        /organism="Triticum aestivum"
        /mol_type="mRNA"
        /cultivar="Chinese Spring"

```

## ORIGIN

```

/db xref="taxon:4565"
/clone="wh3h15"
/feature_type="root"
/dev_stage="feekes" scale 1"
/clone_lib="Y. Ogihara unpublished cDNA library, wh_r"

Query Match      37.5%; Score 546.4; DB 3; Length 704;
Best Local Similarity 88.3%; Pred. No. 3.3e-157;
Matches 606; Conservative 0; Mismatches 76; Indels 4; Gaps 1;

Qy      | 1 | ATTGAACGTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGACGGATGCTTCAT 60
Db      | 23 | ATTGAACGTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGACGGCG---GCAAC 78
Qy      | 61 | CTGTGGCGAGTGGCGGACGGGTGAATATGATGCAACGTAATCAAGAGGGGGGTA 120
Db      | 79 | CTGGCGGCGAGTGGCGGACGGGTGAATATATATGGAACCAAGAGTGGGGGATA 138
Qy      | 121 | ACCGATCGAAGATGTCTAATACCGCATATCTTAAGAGAGAAAGCAGGGGATCGAA 180
Db      | 139 | ACGTACGAAAGTATACGCTAATACCGCATACATCTTAAGATGAAGATGGGGATCGAA 198
Qy      | 181 | GACCTTGGCTTTGGAGCGCCGATGTCTAGTTAGTGTGTGGGTAAAGGCTTAC 240
Db      | 199 | GACCTATCTCTTGGAGCGCCGATATCTGATTAAGTGTGTGTAAAGGCTTAC 258
Qy      | 241 | CAAGCGACGATCAGTATGTTGTCTGAGAGACGACACCACTGGGATGAGACAG 300
Db      | 259 | CAAGCGACGATCAGTATGTTGTCTGAGAGACGACACCACTGGGATGAGACAG 318
Qy      | 301 | GCCGACCTCTACCGGAGGCGACGATGGGGAATTTGACATGGGCCCAAGCCTGATC 360
Db      | 319 | GCCGACCTCTACCGGAGGCGACGATGGGGAATTTGACATGGGCCCAAGCCTGATC 378
Qy      | 361 | CAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGAAGA 420
Db      | 379 | CAGCAATGCGCGGTGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGAAGA 438
Qy      | 421 | AAGGTTAAGCTTAATATCTGACTCATGACGTAATGCAAGAGAGACCGGCTAAC 480
Db      | 439 | AAGGTTAAGCTTAATATCTGACTCATGACGTAATGCAAGAGAGACCGGCTAAC 498
Qy      | 481 | TACGTCCAGACAGCCCGGTAAATAGTAGGTGCAAGCGTTATCGGAATTAATGAGCGGT 540
Db      | 499 | TACGTCCAGACAGCCCGGTAAATAGTAGGTGCAAGCGTTATCGGAATTAATGAGCGGT 558
Qy      | 541 | AAGGTTGCGCGAGCGGCTTTGTAAGTCAGATGTAATCCCGGCTTAACCTGGGAAT 600
Db      | 559 | AAGGTTGCGCGAGCGGCTTTGTAAGTCAGATGTAATCCCGGCTTAACCTGGGAAT 618
Qy      | 601 | TGGCTTGAATACATCAAGGCTAGATGTGGCAGAGGAGGTGAATTCATGTGTAGAG 660
Db      | 619 | TGGCTTGAATACATCAAGGCTAGATGTGGCAGAGGAGGTGAATTCATGTGTAGAG 678
Qy      | 661 | TGAATGCGTGAAGATATGGAAGAC 686
Db      | 679 | TGAATGCGTGAAGATATGGAAGAC 704

RESULT 34
LOCUS      CV292617              875 bp      mRNA      linear      EST 08-JUN-2005
DEFINITION cvf01-2me4-801 Aof01 Asparagus officinalis cDNA clone
ACCESSION  CV292617
VERSION     CV292617.1  GI:52580066
KEYWORDS    Asparagus officinalis
SOURCE      Asparagus officinalis
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
            Asparagaceae; Asparagus.

```

REFERENCE 1 (bases 1 to 875)  
 AUTHORS depamphilis,C., Carlson,J., Ma,H., Soltis,D., Soltis,P.,  
 Openheimer,D., Frohlich,M., Doyle,J., Tankley,S., Webb,M.,  
 Leebens-Mack,J., Landherr,U., Ilut,D. and Wall,K.  
 TITLE Generation of ESTs from early male inflorescences of *Asparagus  
 officinalis*  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Claude depamphilis or James Leebens-Mack  
 Mueller Laboratory  
 Penn State University  
 208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn  
 State University, University Park, PA 16802, USA  
 Tel: 814 863 6413  
 Fax: 814 865 9131  
 Email: cwsd@psu.edu or jhl10@psu.edu  
 The sequence provided is trimmed of vector and low quality regions.  
 Full sequence and original trace file are available from the Plant  
 Genome Network website (<http://pgn.cornell.edu>)  
 Place: aof01-2ms4 row: a column: 01  
 Seq primer: M13F.

FEATURES  
 source  
 1..875  
 /organism="Asparagus officinalis"  
 /mol\_type="mRNA"  
 /db\_xref="PGN:aof01-2ms4-a01"  
 /db\_xref="taxon:4686"  
 /clone="aof01-2ms4-a01"  
 /issue\_type="male inflorescences"  
 /lab\_host="SOLR"  
 /clone\_lib="Aof01"  
 /note="Vector: pBluescript SK (+/-); Site 1: EcoRI;  
 Site 2: XhoI. This is a directionally cloned,  
 non-normalized library. This library has been generated by  
 the Floral Genome Project (FGP). The Floral Genome Project  
 is funded by NSF's Plant Genome Research Program  
 (DBI-0115684). More information about the project can be  
 obtained at <http://fgp.bio.psu.edu>"

ORIGIN

Query Match 37.2%; Score 542.2; DB 7; Length 875;  
 Best Local Similarity 78.6%; Pred. No. 76-156;  
 Matches 686; Conservative 0; Mismatches 183; Indels 4; Gaps 3;

44 GTCGACGACGCGCGGTATAGTGGCGTGAAGGCTTAATGGAATTACTGGCGCTAA 543  
 3 GTGCCAGACGCGCGGTAAAGACAGAGATGCAAGGCTTATCCGAAATGTTGGCGTAA 62  
 544 GGGTCGCGAGCGCGCTTGTATAGTCAGATGTGAATCCCGGGCTTAACCTGGAAATGC 603  
 63 GCGTCTGTAAGTGGCTTTTCAAGTCCGCCGTCAATCCAGGGCTCAACCTGGACAGCG 122  
 604 GTTTGAAACTACAGAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTAGACGTGA 663  
 123 GGTGGAATCTACCAAGCTGAGATACGATGAGGGGAGAGGAATTCGGGTGAGCGGTGA 182  
 664 AATGGTGAAGATATAGGAAGACATCGATGAGGAGCGAGCCTCGGGTTAACTACACTAC 723  
 183 AATGCGTAGAGATCGGAGAAACACCAACGCGGAAGCACTGTGCTGGCGGACACTGAC 242  
 724 GCTCATGACAGAAACCTGGGAGGAGCAACAGATTAGATACCTGTGATCCAGCCCTTA 783  
 243 ACTGAGAACGCAAACTAGGGAGGAGCAAAATGGATTAGATACCCAGTAGTCTTACCCGTA 302  
 784 AACGATGTCAACTAGTGTGTGGGCTTATTAGGCTTG--GTACGAAGCTAAACGCGTGA 841  
 303 AACGATGTACTAGTACGCGCTGTGCGATGACCCGCGAGTCTGTAGCTAACGCGTTAA 362  
 842 GTTGACCCGCTGGGAGATACGCTCGCAAGATTAAACTCAAGGAATTGACGGGAGCCCG 901  
 363 GTATCCCGCTGGGAGATACGTTCCCAAGAAATCAAAAGGAATTGACGGGAGCCCG 422  
 902 CACAAGCGGTGATTAATGTGATTAATTCGATGACGGAAGAAACCTTACCTACCCCTTG 961

Db 423 CACAAGCGGTGAGCATGTGTTTAATTCATGATCAAGCAAGAACTTACAGGCGCTTG 482  
 QY 962 ACATGTAGCAATTTTCTAGAGATTAGTATG--CTTGGGAAAGCTAACACAGGTGCTG 1020  
 Db 483 ACATGCCCGAATCTCTTGAAGAGAGGGGTCTTGGGAAAGCGGACACAGGTGGTG 542  
 QY 1021 CATGCTCTCGTCAAGCTCGTGTGAGATGTGGGTTAAGTCCCGCAACGAGCGCAAC 1080  
 Db 543 CATGCTCTCGTCAAGCTCGTGTGAGATGTGGGTTAAGTCCCGCAACGAGCGCAAC 602  
 QY 1081 CTGTCTATTAAATGCGATCATTTGGTTGGGCACTTAATAGACTGCGCGTGAACAACG 1140  
 Db 603 CTGCTGTTAATGTGCAACCATGATGATGGAACCTGGAACAGACTCCGCTGATAGCG 662  
 QY 1141 GAGGAAGTGGGAGATGAGTCAAGTCCATAGGCGCTTATGGGTAGGGCTTCACAGTAA 1200  
 Db 663 GAGGAAGTGGGAGATGAGTCAAGTCCATAGGCGCTTATGGGTAGGGCTTCACAGTAA 722  
 QY 1201 TACAATGCGCGTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260  
 Db 723 TACAATGCGCGGAGCAAAAGGCTCGGATCCGCGAGGCTGAGCTAACCAAAACCGT 782  
 QY 1261 TCGTAGTCCGATCGGATCGGATCTGCAACTCGCTGGAATCGGATGCTAGTATGCG 1320  
 Db 783 CCTCAGTTCGGAATGCAAGCTGCAACTCGCTGCATGAAGCGGATGCTAGTATGCG 842  
 QY 1321 GGATCAG--CATGTCGCGGTGAATACGTTCCCGG 1352  
 Db 843 CGGTACGCAATACGCGGATGAATTCGTTCCCGG 875

RESULT 35  
 BZ502068/c  
 LOCUS BZ502068 1096 bp DNA linear GSS 16-DEC-2002  
 DEFINITION BON1125R BO.1.6.2\_KB.tot Brassica oleracea genomic clone BON1125,  
 genomic survey sequence.  
 ACCESSION BZ502068  
 VERSION BZ502068.1 GI:27019377  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 1096)  
 Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
 Utecherack,T.R., Wortman,J.R., White,O.R. and Town,C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 15805490  
 Other\_GSSs: BON1125TF  
 CONTACT Contact: Chris Town  
 TITER 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@ciagr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.

FEATURES  
 source  
 1..1096  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="T01000D3"  
 /db\_xref="taxon:3712"  
 /clone="BON1125"  
 /clone\_lib="BO\_1.6.2\_KB.tot"  
 /note="Vector: pBOS1; Site 1: BstXI, 1.6-2 kb sheared  
 total DNA inserted into pBOS1 using BstXI linkers"

ORIGIN



QY 409 CAGTGGAGAGAAAAGCTTACGGTAAATATCTGACTCATGACGGTATCGACAGAA 468  
Db 449 TATCAGGAAGAAAGAAATTAATATATATATTAATTAATGACGGTACCTGAAAGAAATA 508  
QY 469 GCACCGGGCTAATACGATGCGCAGACCGCGGTAAATACGTAAGGTCGAACGTTAATCGGA 528  
Db 509 GCACCGGGCTAATACGATGCGCAGACCGCGGTAAATACGTAAGGTCGAACGTTAATCGGA 568  
QY 529 ATTACTGGCGGTAAAGGGTCCGACGCGGCTTTGTAAAGTCAGATGTGAATCCCGGGCT 588  
Db 569 ATTATTTGGCGGTAAAGCGTACGTAAGCGGGTTTGTAAAGCAGATGTGAATCCCTGGGCT 628  
QY 589 TAACCTGGGAATTCGCTTTGAAACTCAAGGCTAAGATGTGCGACAGAGGAGGTGAATTC 648  
Db 629 CAACCTTGGGAATTCGCTTTGAAACTCAAGGCTAAGATGTGCGACAGAGGAGGTGAATTC 688  
QY 649 CATGTGTGAGCAATGAAATGCGTAAAGATATGGAAGAACATCGATGCGGACGACGCTCC 708  
Db 689 CACATGTAGCAGTGAATGCGTAAATGATGTGGAAGAAATCCGATGGCGAAGGCAAGTCTCC 748  
QY 709 TGGGTTAACTGACGCTCATGACGAAAGCGTGGGGAG 747  
Db 749 TGGGATTAACACTGACGCTCATGATGAGAAAGCGTGGGGAG 787

RESULT 37  
BH527452 862 bp DNA linear GSS 13-DEC-2001  
LOCUS BOGVP3CTR BOGZ Brassica oleracea genomic clone BOGVP36, genomic  
DEFINITION survey sequence.  
ACCESSION BH527452  
VERSION BH527452.1 GI:17735537  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 862)  
Ayele, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Uteirbeck, T. R., Wortman, J. R., White, O. R. and Town, C. D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490  
JOURNAL PUBMED  
COMMENT Other GSSs: BOGVP36TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Classes: sheared ends.  
FEATURES  
source  
1..862  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOGVP36"  
/clone\_1kb="BOGV"  
/note="Vector: pHD01, Site 1: BactX; 2-3 kb sheared  
genomic DNA inserted into pHD01 using BactX linkers"

ORIGIN  
Query Match 36.1%; Score 526.4; DB 9; Length 862;  
Best Local Similarity 77.7%; Pred. No. 5,4e-151;  
Matches 685; Conservative 0; Mismatches 171; Indels 26; Gaps 3;  
QY 207 GTCTATTAAGCTAGTTGGTGGGTAAAGCCCTACCAAGCGACGATCACTAGTGTCTG 266

Db 1 GTCTATTAAGCTAGTTGGTGGGTAAAGCCCTACCAAGCGACGATCACTAGTGTCTG 60  
QY 267 AGAGACGACCAACCACTGGGACTGAGACACGGCCCAAGCTCTTACGGAGCAGAG 326  
Db 61 AGAGATGATACGACCACTGGGACTGAGACACGGCCCAAGCTCTTACGGAGCAGAG 120  
QY 327 TGGGGAATTTTGGCAATGGGCGCAAGCTGATCAGCAATGGCGCGTGAATGAAGAG 386  
Db 121 TGGGGAATTTTCCGCAATGGGCGCAAGCTGATCAGCAATGGCGCGTGAATGAAGAG 180  
QY 387 CTTTCGGGTTGTAAGCTCTTTCAGTCGAGAAAGAAAGTTACGATTAATGCTGACT 446  
Db 181 CTTTCGGGTTGTAAGCTCTTTCAGTCGAGAAAGAAAGTTACGATTAATGCTGACT 217  
QY 447 CATGACGATTCGACAGAAAGACACGGCTTAATCTGTCGACAGACCGCGGTAAATTCG 506  
Db 218 AATGACGATTCGAGGGAATTAAGCATCGCTTAATCTGTCGACAGACCGCGGTAAATTCG 277  
QY 507 TAGGTCGAACCTTAATCGGAATTAATCGGCGTAAAGGGTGGCGGCGGCTTTGTAA 566  
Db 278 GAGATGCAACGCTTATCCGAAATGATTGGCGTAAAGCGTCTGTAGTGGCTTTTAA 337  
QY 567 TCAATGTGAATCCCGGCGCTTAACCTGGAATTCGCTTTGAAACTCAAGGCTAGAGT 626  
Db 338 TCCGCGCTCAATCCAGGCGCTCAACCTGACAGGCGGTGAATCAAGCTTGAAGT 397  
QY 627 GTGCGAGGAGGAGGTGAATTCATGTGTACAGTGAATTCGTAAGATATGGAAGAC 686  
Db 398 ACGGATGGGGGACAGAGGAATTTCCGTTGACCGGTGAATTCGTAAGATGGAAGAAC 457  
QY 687 ATCGATGCGGAAGGACGCTCTGGGTTAACATGACGCTCATCAAGAAAGCGTGGGA 746  
Db 458 ACCAAGCGCGAAGCACTCTGCTGGCCGACATGACATGACAGACGAAAGCTTAAGGGA 517  
QY 747 GCAACAGGATTAAGTACCTGTAGTCCAGCCCTTAACCATGTCAATGATTGTTGG 806  
Db 518 GCGAATGGGATTAAGTATCCCAAGTATCTGACCGTTAAGATGATCTAGGGCGTGTG 577  
QY 807 CTTTATTAGGCTTG--GTAAAGAGCTTAACCGCTGAAGTTGACCGCTGGGGAAGTACGCT 864  
Db 578 CGTATCGACCGGTGAGTGTGTAGCTTAACCGCTGAAGTTGACCGCTGGGGAAGTACGCT 637  
QY 865 CGCAAGATTAAGTAACTCAAGGAATTTGACGGGACCGGACAGGCGGTGATTTATGTGAT 924  
Db 638 CGCAAGATTAAGTAACTCAAGGAATTTGACGGGACCGGACAGGCGGTGATTTATGTGAT 697  
QY 925 TAATTCGATGCAAGCGCAAAACCTTAACCTTACCATGTAGCGAATTTCTTACAGA 984  
Db 698 TAATTCGATGCAAGCGCAAAACCTTACCATGTAGCGAATTTCTTACAGA 757  
QY 985 TAGATTAGTG--CTTGGGAAACGTTAACACAGGTGCTGATGCTGCTACGCTGTGTC 1043  
Db 758 GAGAGGGGTGCTTTCGGGAAAGCGCGACACAGGTGTGATGCTGTGCTGCTGTGCC 817  
QY 1044 GTGAGATGTTGGTTAAGTCCCGCAAGGCGCAACCTTGT 1085  
Db 818 GTAAGTGTGGTTAAGTCCCGCAAGGCGCAACCTTGT 859

RESULT 38  
BH562180/c 846 bp DNA linear GSS 14-DEC-2001  
LOCUS BOGZN83TF BOGZ Brassica oleracea genomic clone BOGZN83, genomic  
DEFINITION survey sequence.  
ACCESSION BH562180  
VERSION BH562180.1 GI:17814020  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 846)  
 AUTHORS Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
 Utecherback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
 TITLE Whole genome shotgun sequencing of *Brassica oleracea* and its  
 application to gene discovery and annotation in Arabidopsis  
 JOURNAL Genome Res. 15 (4), 487-495 (2005)  
 PUBMED 15805490  
 COMMENT Other GSSs: BOG2N83TR  
 Contact: Chris Town  
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TP  
 Class: sheared ends.

FEATURES  
 source  
 1..846  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000DH3"  
 /db\_xref="taxon:3712"  
 /clone\_1lb="BOG2N83"  
 /note="Vector: pHO51; Site 1: Bactx1; 2-3 kb sheared  
 genomic DNA inserted into pHO51 using Bactx1 linkers"

ORIGIN  
 Query Match 36.1%; Score 526.2; DB 9; Length 846;  
 Best Local Similarity 78.0%; Pred. No. 6.2e-151;  
 Matches 659; Conservative 0; Mismatches 183; Indels 3; Gaps 2;

487 CCAGCAGCCCGGCTATAGTGGGCGTGAAGCTTATCGGAGCTTAAGG 546  
 846 CCAGCAGCCCGGCTATAGTGGGCGTGAAGCTTATCGGAGCTTAAGG 787  
 547 TGCGCAGCGCGCTTGTAGTGCATGTGAATCCCGGCGCTTAACCTGGGATTCGCT 606  
 786 TCTGTAGTGGCTTTTAACTCCGCGCTCAATCAGGCTCAACCTGGGATTCGCT 727  
 607 TGAATCAAGGCTAGTGTGGGCGTGAAGCTTATCGGAGCTTATCGGAGCTTAAGG 666  
 726 GGAATCAAGGCTAGTGTGGGCGTGAAGCTTATCGGAGCTTATCGGAGCTTAAGG 667  
 667 GCGTATAGTATGGAAGATGATGATGCGGAGCGGAGCTTATCGGAGCTTATCGGAG 726  
 666 GCGTATAGTATGGAAGATGATGATGCGGAGCGGAGCTTATCGGAGCTTATCGGAG 607  
 727 CATGCAAGGCTAGTGTGGGCGTGAAGCTTATCGGAGCTTATCGGAGCTTATCGGAG 786  
 606 GATGCAAGGCTAGTGTGGGCGTGAAGCTTATCGGAGCTTATCGGAGCTTATCGGAG 547  
 787 GATGCAAGGCTAGTGTGGGCGTGAAGCTTATCGGAGCTTATCGGAGCTTATCGGAG 844  
 546 GATGCAAGGCTAGTGTGGGCGTGAAGCTTATCGGAGCTTATCGGAGCTTATCGGAG 487  
 845 GATGCAAGGCTAGTGTGGGCGTGAAGCTTATCGGAGCTTATCGGAGCTTATCGGAG 904  
 486 TCCGCGCTGGGAGTACGTTCCGCAAGATGAATCTCAAGGAAATTGAACGGGCGCGCAC 427  
 905 AAGGCGTGTGATATGATGATTAATGATGCAACGGGAAACCTTACCTACCTTACCA 964  
 426 AAGGCGTGTGATATGATGATTAATGATGCAACGGGAAACCTTACCTACCTTACCA 367  
 965 TGTAGCAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTTACCAAGCTTCGAT 1023  
 366 TGTAGCAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTTACCAAGCTTCGAT 307  
 1024 GGTGTGTGTGAGT 1083  
 306 GGTGTGTGTGAGT 247

1084 GTCAATTAATGCGCATATGTTGTTGGGCACTTTAATGAGACTGCGGTGACAAACCGGAG 1143  
 246 GTGTTTACTTCCACCGTTGAGTTTGAACCCCTBACAGACTGCGGTGATTAACCGGAG 187  
 1144 GAAAGGCGGATGAGCGTCAAGTCTTATGAGGCTTATGAGGCTTACACGTAATAC 1203  
 186 GAAAGGCGGATGAGCGTCAAGTCTTATGAGGCTTATGAGGCTTACACGTAATAC 127  
 1204 AATGCGCGGTACAGAGGTTTCCAAACCGGAGGCGGAGCTTATGAGGAGGCGGTG 1263  
 126 AATGCGCGGTACAGAGGTTTCCAAACCGGAGGCGGAGCTTATGAGGAGGCGGTG 67  
 1264 TAGTCCGATGAGGATGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 1323  
 66 CAGTTGCGATGAGGATGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCGG 7

QY 1324 TCAGC 1328  
 DB 6 TCAGC 2

RESULT 39  
 BH705306 873 bp DNA linear GSS 20-FEB-2002  
 BOMIV08TF BO\_2\_3 KB Brassica oleracea genomic clone BOMIV08,  
 LOCUS genomic survey sequence.  
 DEFINITION  
 ACCESSION BH705306  
 VERSION BH705306.1 GI:18787817  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 873)  
 Utecherback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
 TITLE Whole genome shotgun sequencing of *Brassica oleracea* and its  
 application to gene discovery and annotation in Arabidopsis  
 JOURNAL Genome Res. 15 (4), 487-495 (2005)  
 PUBMED 15805490  
 COMMENT Other GSSs: BOMIV08TR  
 Contact: Chris Town  
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TP  
 Class: sheared ends.

FEATURES  
 source  
 1..873  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000DH3"  
 /db\_xref="taxon:3712"  
 /clone\_1lb="BO\_2\_3 KB"  
 /note="Vector: pHO51; Site 1: Bactx1; 2-3 kb sheared  
 genomic DNA inserted into pHO51 using Bactx1 linkers"

ORIGIN  
 Query Match 36.0%; Score 524.8; DB 9; Length 873;  
 Best Local Similarity 77.6%; Pred. No. 1.7e-150;  
 Matches 660; Conservative 0; Mismatches 187; Indels 3; Gaps 2;

448 ATGACGATATGACAGAGAGACCGGCTTAATGATGCAAGCTTATCGGAGCTTATCGT 507  
 DB 858 ATGACGATATGACAGAGAGACCGGCTTAATGATGCAAGCTTATCGGAGCTTATCGT 799  
 QY 508 AGGCTCAAGCTTATGCAATTAATGATGCAAGCTTATGCAAGCTTATGCAAGCTTAT 567



Db 798 AGGATGCAAGCGTATCCGGATGATGGGGCTTAAGCGTCTGTAGTGGCTTTTAAAGT 739  
Qy 568 CAGATGTGAATCCCGGGCTTAACTGGGAAATGGCTTTGAATACAGGCTTGAGTG 627  
Db 738 CCGCGGTAAATCCAGGGCTCAACCTGGACAGGGGTGAAGAACTACCAACTTGAGTA 679  
Qy 628 TGGCAGAGGAGGTGAATTCATGTGTAGAGTGAATGCGTAGAGTATGGAAGACA 687  
Db 678 CCGTAGGGGACAGGAAATTCGGGTGAGCGGTGAATGCTTAGAGTCGAAAGAACCA 619  
Qy 688 TCGATGGCGAAGGACGCTCTGGGTTAACCTGACGCTCATGACAGAAAGCGTGGAG 747  
Db 618 CCAAGCGGAAAGCACTGCTGCTGGGCGGACATGACACTGAGAGAGAAAGCTAGGGAG 559  
Qy 748 CAAACGATTTAGATACCTGTGATGCCAGCGCTTAAACGATGTCACTAGTTGGGC 807  
Db 558 CGAATGGGATTTAGATACCCAGTAGCTCTAGCCGTAACGATGATACAGCGCTGTGC 499  
Qy 808 CTTATTAGGCTTG--GTAAAGGCTAACGGGTGAAGTTGACCGCTGGGAGTACGGTC 865  
Db 498 GTATGACCCCGTGCAGTGTCTAGCTTAACGCTTAAGTATCCCGCTGGGAGTACGTT 439  
Qy 866 GCAAGATTAAACTCAAGAAATGACGGGACCGGACAAAGCGGTGATTTATGTGAT 925  
Db 438 GCAAGAAATGAATCTCAAGAAATGACGGGCGCGGACAAAGCGGTGAGCATGTGTT 379  
Qy 926 AATTGATGCAACGCGAAACCTTACCTACCTTGAATGTAGCGAATTTTCTAGAGAT 985  
Db 378 AATTGATGCAAGCGAAGAACCTTACCGGCTTGAATGTAGCGAATTTTCTAGAA 319  
Qy 966 AGATTAGTG--CTTGGGAGCGCTAACAGAGTGTGACGATGCTGTGCTGCTGCTG 1044  
Db 318 AGAGGGGTGCTTGGGAAACGCGAACAGAGTGTGATGCTGTGCTGCTGCTGCTG 259  
Qy 1045 TGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCTTGTATTAATTTCCATGTT 1104  
Db 258 TAAAGTGTGGGTAAAGTCCCGCAACGAGCGCAACCTTGTATTAATTTCCATGTT 199  
Qy 1105 GTTGGGCACTTTAATGAGCTGCGCGTGAACAAACCGAGAGAAAGTGGGATGACTCA 1164  
Db 198 GTTGGAAACCTGAAACAGACTGCGGTGATTAAGCGGAGAAAGTGAAGATACCTCA 139  
Qy 1165 TCCCTATGCGCTTATGCTAGGCTTCAACGATTAATCAATGGGCGCTGACGAGGTTG 1224  
Db 138 TCATATGCGCTTATGCTAGGCTTCAACGATTAATGAGCGGAGCAAAAGGCTG 79  
Qy 1225 CCAACCGCGAGAGGAGCTAATCTCAGAAAGCGGTGATGCTGAGTCTGCA 1284  
Db 78 CGATCCCGGAGGAGCTAATCTCAGAAACCGGTCTCACTTGGATGCAACTGCA 19  
Qy 1285 ACTGCACTCC 1294  
Db 18 ACTGCGCTGC 9

RESULT 40  
BH547523/c 875 bp DNA linear GSS 14-DEC-2001  
LOCUS BOGH290TF BOGH Brassica oleracea genomic clone BOGH290, genomic  
DEFINITION survey sequence.  
ACCESSION BH547523  
VERSION BH547523.1 GI:17799304  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta, eudicotyledons, core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 875)  
Ayele, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Uteirbeck, T. R., Mortman, J. R., White, O. R. and Town, C. D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis

JOURNAL Genome Res. 15 (4), 487-495 (2005)  
PUBMED 15805490  
COMMENT Other\_GSSs: BOGH290TR  
Contact: Chris Town  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.  
FEATURES  
source location/Qualifiers  
1..875  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="T01000D43"  
/db\_xref="taxon:3712"  
/clone="BOGH290"  
/clone\_lib="BOGH"  
/note="Vector: pHD01, Site 1: BexXI, 2-3 kb sheared  
genomic DNA inserted into pHD01 using BexXI linkers"

ORIGIN  
Query Match 36.0%; Score 524.6; DB 9; Length 875;  
Best Local Similarity 77.6%; Pred. No. 2e-150;  
Matches 696; Conservative 0; Mismatches 174; Indels 27; Gaps 4;  
Qy 235 GCTTACCAAGGCGAGATCACTAGTGTGTGAGAGACGACCACTGGGACTGA 294  
Db 873 GCTACCCAAAGCGATGATCAGTACGTGGTCCGAGAGATGATCAGCACACTGGGACTGA 814  
Qy 295 GACACGCGCCGACACTCTTACCGGAGCAGACGATGGGGAATTTTGACAAATGGGCGAAC 354  
Db 813 GACACGCGCCGACACTCTTACCGGAGCAGACGATGGGGAATTTTGACAAATGGGCGAAC 754  
Qy 355 CTGATCCAGCAATGCGCGGTGAGTGAAGAAAGCCCTTGGGTTGTAAGCTCTTCACTG 414  
Db 753 CTGACGAGCAATGCGCGGTGAGTGAAGAAAGCCCTTGGGTTGTAAGCTCTTCACTG 695  
Qy 415 AGAAGAAAGGTTACGGTAAATATCGTGACTCATGACGATGACAGAAAGACCG 474  
Db 694 AGAAG-----ACCATGACGGTATCTGGGGAATTAAGCATCG 658  
Qy 475 GCTAATCACTGCTCAAGCAGCGCGCTAATTAAGTGGTGCAGCGCTTAATGGAATTA 534  
Db 657 GCTAATCTGTGCAAGCAGCGCGCTAATTAAGTGGTGCAGCGCTTAATGGAATTA 598  
Qy 535 GGGGCTAAAGGATCGCAGCGGCTTGTATGATGATGTAATCCCGGGCTTAACCT 594  
Db 597 GGGGCTAAAGGATCGTGAAGGCTTGTATGATGATGTAATCCCGGGCTTAACCT 538  
Qy 595 GGGGATTCGTTTAAACTCAAGGCTAAGTGTGAGCAGAGGAGTGAATTCATGTC 654  
Db 537 TGAACGCGGTGAAACTTACCAAGCTTGAATGATGAGGAGGAGAAATTTCCGGTG 478  
Qy 655 TAGCAGTGAATGCTGATGATGAGTGAAGAAACATCGATGCGAAGCGCTCTGGGTT 714  
Db 477 GAGGCGTGAATGCTGATGATGAGTGAAGAAACACCAACCGGGAAGCACTGTGCTGGCC 418  
Qy 715 AACACTGACGCTCAATGACGAAAGCGTGGGAGCAACAGATTAGTACCTGGTATGTC 774  
Db 417 GACACTGACCTGAGAGACGAAACGTAGGGAGCGAATGGGATTAATACCCAGTATGC 358  
Qy 775 CACGCCCTAAAGATGTCATAGTGTGCGCTTATTAGCTTG--GTAAACGAAGTGA 832  
Db 357 CTACCCGTAAGCAATGATATTAAGCGCTGTGCTATGACACCGTGAAGTCTGTAGCTA 298  
Qy 833 ACGCGTGAATTTGACCGCGCTGGGAGTACGCTGCCAAGATTAAATCTCAAGAAATTTGAC 892  
Db 297 ACGGTTTAAGTATCCCGCTGGGAGTACGTTCCCAAGAAATGAATCTCAAGAAATTTGAC 238  
Qy 893 GGGGACCGGCAACAGCGGTGATATGATTAATTCATGACACGGAAGAAACCTTAC 952



Db 237 GGGGGCCCGCAGACCGGTGGAGCATGTGTTTAATTCATGCAAGCAGCAACCTTAC 178  
Qy 953 CTACCTTGACATGAGGAATTTCTAGATAGATTAGTG-CTTCGGGAGCGTAACA 1011  
Db 177 CAGGGCTTGACATGCGCGCAATCTCTTGAAAGAGGGGTCTTCGGGAGCGGACA 118  
Qy 1012 CAGGTGCTGATGAGTGTGTCAGCTCGTGTGATGATGTTGGTTAAGTCCGCAAG 1071  
Db 117 CAGGTGCTGATGAGTGTGTCAGCTCGTGTGATGATGTTGGTTAAGTCCGCAAG 58  
Qy 1072 AGCGCAACCTTGTCATTATTCATCAATCTTTGGTTGGCACTTATATAGACTGC 1128  
Db 57 AGCGCAACCTTGTCATTATTCATCAATCTTTGGTTGGCACTTATATAGACTGC 1  
RESULT 41  
BZ426275 889 bp DNA linear GSS 13-DEC-2002  
LOCUS BOOAB1TR.B0.1.6.2 KB tot Brassica oleracea genomic clone BOOAB1,  
DEFINITION genomic survey sequence.  
ACCESSION BZ426275  
VERSION BZ426275.1 GI:26667787  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 889)  
AUTHORS Ayle, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Ullrich, T. R., Mortman, J. R., White, O. R., and Town, C. D.  
TITLE Whole genome shotgun sequencing of Brassica oleracea and its  
JOURNAL application to gene discovery and annotation in Arabidopsis  
PUBMED Genome Res. 15 (4), 487-495 (2005)  
COMMENT 15805490  
Other GSSs: BOOAB1TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seg primer: TR  
Class: sheared ends.  
Location/Qualifiers  
1. 889  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOOAB1"  
/clone\_1bp="BO.1.6.2 KB tot"  
/note="Vector: pHD01; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pHD01 using BstXI linkers"  
ORIGIN  
Query Match 36.0%; Score 524.4; DB 9; Length 889;  
Best Local Similarity 77.4%; Pred. No. 2.3e-150;  
Matches 697; Conservative 0; Mismatches 176; Indels 27; Gaps 4;  
Qy 186 TGGCCTTTGAGCGCGCCAGTCTGATTAGTACTAGTGTGGGTAAAGGCTTACCAAG 245  
Db 876 TCGCCCGCAGAGGAGGCTCGCGCTGTGATGCTAG-TGGTAGCAAAATAGCTTACCAAG 818  
Qy 246 CGACGATCACTAGTGTGCTGAGAGCAACCAACCACTGGAGCTGAGCAACGAGCCCA 305  
Db 817 CGAGATCACTAGTGTGCTGAGAGCAACCAACCACTGGAGCTGAGCAACGAGCCCA 758  
Qy 306 GACTTCCTAGCGGAGCAACCACTGGGAAATTTGACCAATGGGCGCAAGCCTGATCCGCA 365  
Db 757 GACTTCCTAGCGGAGCAACCACTGGGAAATTTGCGCAATGGGCGCAAGCCTGAGCGAGCA 698

Qy 366 ATGCCGCTGAGTGAAGAAGGCTTCGGGTTGTAAAGCTTTCACTGACGAGAGAAAAG 425  
Db 697 ATGCCGCTGAGTGAAGAAGGCTTACGGGTCTGTAATCTTTTCCAGAGAGA----- 643  
Qy 426 TTACGTAATATATCTGACTCATGACGGTATCGACAGAGAGAGCAAGGCTTAACACTAGT 485  
Db 642 -----AGCAATGACGGTATCTGGGGAATTAAGATAGGCTTAACCTGT 601  
Qy 486 GCCAGACCGCGGTAAATAGCTAGGGTCAAGCGTTAATCGGAATTACTGGGCGTTAAG 545  
Db 600 GCCAGACCGCGGTAAATAGCTAGGGTCAAGCGTTAATCGGAATTACTGGGCGTTAAG 541  
Qy 546 GTGCCAGCGGCTTTGTAATGTCAGATGTAATCCCGGGCTTAACCTGGGAATTCGCT 605  
Db 540 GTCTGTAGTGTCTTTTAAATGTCGCGCTCAATCCAGGGCTCAACCTGACAGGCG 481  
Qy 606 TTGAATTAACAAGGTAGATGTGGCAGAGGAGGTGAATTCATGTGTAGCACTGA 665  
Db 480 TGAATTAACAAGGTAGATGTGGCAGAGGAGGTGAATTCATGTGTAGCACTGA 421  
Qy 666 TCCGTAGATATGAGAGAAACATGATGCGAGCGACGCTCTGGGTTAACAATGACGC 725  
Db 420 TCGTATGATGAGAGAAACATGATGCGAGCGACGCTCTGGGTTAACAATGACGC 361  
Qy 726 TCAATGAGAAAGCGTGGGAGCAACAGATTAATACCTGATGATGATGATGATGAT 785  
Db 360 TGAAGAGCAAAAGCTAGAGGAGCGAATGATTAATACCTGATGATGATGATGAT 301  
Qy 786 CGATGCACTAGTGTGGGCTTATTAAGCTTG--GTAAAGCACTAACCGCTGAGT 843  
Db 300 CGATGCACTAGTGTGGGCTTATTAAGCTTG--GTAAAGCACTAACCGCTGAGT 241  
Qy 844 TCACCGCTGAGAGATACGCTGCAAGATTAATCAAGAAATGACGGGAGCCGCA 903  
Db 240 ATCCGCGCTGGGAGATAGCTTGCAGAAATGAACTCAAGAGAAATGACGGGAGCCGCA 181  
Qy 904 CAAAGCGTGTGATGATGATTAATTCATGCAAGCGGAAACCTTACAGGGCTTGAC 963  
Db 180 CAAAGCGTGTGATGATGATTAATTCATGCAAGCGGAAACCTTACAGGGCTTGAC 121  
Qy 964 ATGTAGCGAATTTCTAGAGATAGATTAGTG-CTTCGGGAGCGCTTACAGGTGCTGCA 1022  
Db 120 ATGTAGCGAATTTCTAGAGATAGATTAGTG-CTTCGGGAGCGCTTACAGGTGCTGCA 61  
Qy 1023 TGGCTGTGCTGAGCTGTGCTGAGATGTTGGTTAAGTCCCGAAGAGCGCAACCTT 1082  
Db 60 TGGCTGTGCTGAGCTGTGCTGAGATGTTGGTTAAGTCCCGAAGAGCGCAACCTT 1  
RESULT 42  
CF566443 745 bp mRNA linear EST 23-SEP-2003  
LOCUS 1115107G11.X1 C. reinhardtii CC-1690 (mt+), Stress III  
DEFINITION (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 3', mRNA  
sequence.  
ACCESSION CF566443  
VERSION CF566443  
KEYWORDS EST.  
SOURCE CF566443.1 GI:3492526  
ORGANISM Chlamydomonas reinhardtii  
Chlamydomonas reinhardtii  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;  
Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.  
REFERENCE 1 (bases 1 to 745)  
AUTHORS Groisman, A., Chang, C.-W., Davies, J., Harris, B., Hauser, C.,  
Merchant, S., Quinn, J., and Shrago, J.  
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
JOURNAL Unicellular System for Analyzing Gene Function and Regulation in  
COMMENT Vascular Plants. Project: 1115  
Contact: Charles Hauser  
DCMB Box 91000  
Duke University

Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: chausser@duke.edu  
High quality sequence stop: 745.  
Location/Qualifiers

## FEATURES

source

1..745

/organism="Chlamydomonas reinhardtii"

/mol\_type="rRNA"

/db\_xref="taxon:3055"

/clone\_1ib="C. reinhardtii CC-1690 (mt+), Stress III

(normalized), Lambda Zap II"

/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:

XhoI; Stress condition III library, constructed by

Jeanette Quinn and Ching-Wen Chang, combines cDNAs from

CC-1690 cells grown to mid-log phase in copper-free TAP

medium (see Quinn and Merchant (1998) Methods in

Enzymology, 2997:263-279) in a shaking (250 rpm) illuminated

(about 100 micromole/m2/sec) incubator at 22 C (see Quinn

and Merchant (1998) Methods in Enzymology, 2997:263-279);

CC-1690 cells grown to mid-log phase in low Fe (1

micromolar Fe) TAP medium (see La Fontaine S, Quinn JM,

Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J,

Merchant S, LaFontaine et al. (2002) Eukaryotic Cell,

1:736-757) in a shaking illuminated incubator (same

conditions as above). CC-1690 cells were grown to mid-log

phase in TAP medium in a shaking illuminated incubator to

a density of 8x10<sup>6</sup> cells/ml. The flask was transferred to

a shaking platform (200 rpm) at room temperature (23C) 12

micromole/m2/sec illumination and bubbled in a stoppered

flask with 98% nitrogen, 2% CO<sub>2</sub> gas mixture for 1 hour

prior to harvesting for RNA isolation (as per Quinn JM,

Barriaco P, Eriksson M, Merchant S, Quinn et al. (2000) JBC

275:6080-6089); CC-1690 cells grown to mid-log phase

(3x10<sup>6</sup> cells/ml) in TAP medium in a shaking (150 rpm)

illuminated (70 mole photon/m2/sec) incubator at 27 C.

Cells were diluted to 1x10<sup>6</sup> cells/ml, transferred to high

light (11000 mol photon/m2/s) with shaking (150 rpm) and

sampled at (0.5, 1.2, 4.6, 12 hrs). PolyA mRNA was purified

from each sample, pooled and cDNA synthesized (see Shragar

et al, Plant Physiol. 131, 401-408 for details). The cDNA

was directionally cloned into lambda Zap II (Stratagene)

in the EcoRI (5') and XhoI (3') sites. pBluescript II SK-

plasmids were excised from the lambda Zap clones by

superinfection with ExAssist (Stratagene) phage. The

library was normalized using method 4 described in Bonaldi

et al., (1996) Genome Research 6: 791-806."

## ORIGIN

Query Match

Best Local Similarity 84.3%; Score 521.8; DB 6; Length 745;

Matches 622; Conservative 0; Mismatches 113; Indels 3; Gaps 3;

## ORIGIN

468 AGCACCGGCTAAGGCTGCGCAGCGCGGCTTAATAGCTGCAAGCTTAATCGG 527

## ORIGIN

8 AGCACCGGCTAAGGCTGCGCAGCGCGGCTTAATAGCTGCAAGCTTAATCGG 67

## ORIGIN

528 AATTACTGGGCGTAAAGGCTGCGCAGCGCGGCTTTGTAAGTCAGATGTGAATCCCGGCG 587

## ORIGIN

68 AATTACTGGGCGTAAAGGCTGCGCAGCGCGGCTTTGTAAGTCAGATGTGAATCCCGGCG 127

## ORIGIN

588 TTAACCTGGGAATTCGTTGAATCAAGCTAGAGTGGCAGAGGAGAGTGGATT 647

## ORIGIN

128 TCAACCTGGGAATTCGTTGAATCAAGCTAGAGTGGCAGAGGAGAGTGGATT 187

## ORIGIN

648 CCATGTGAGCAGTGAATGCGTAGAGATATGGAAGAAATCATGAGCGCAAGGACCTTC 707

## ORIGIN

1188 CCAGGTGAGCGGTGTAAGTGCATGAGATCTGAGGAATACCGGTGCGAAGCGGCCCC 247

## ORIGIN

708 CTGGGTTAAACATGACGCTCATGACGGAAGCGTGGGAGGAAACAGGATTAGATACCT 767

## ORIGIN

248 CTGACGAGAGACTGACGCTTCAGGTGCGAAGCGTGGGAGGAAACAGGATTAGATACCT 307

## ORIGIN

768 GGTAGTCCACGCGCTTAAAGGCTGCGCAGCGCGGCTTTGTAAGTCAGATGTGAATCCCGGCG 826

## ORIGIN

308 GGTAGTCCACGCGCTTAAAGGCTGCGCAGCGCGGCTTTGTAAGTCAGATGTGAATCCCGGCG 367

## ORIGIN

827 AAGCTAACGCGTGAAGTTGACCGCTTGGGAGTACCGTCCGAAATTAATCAAGAGA 886

## ORIGIN

368 GAGCTAACGCGTGAAGTTGACCGCTTGGGAGTACCGTCCGAAATTAATCAAGAGA 427

## FEATURES

source

1..842

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

## FEATURES

source

1..842

/organism="Brassica oleracea"

## FEATURES

source

1..842

/organism="Brassica oleracea"

## FEATURES

source

1..842

/organism="Brassica oleracea"

## FEATURES

source

1..842

/organism="Brassica oleracea"

```

/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGAK19"
/clone_1ib="BOGAK"
/notes="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOSt using BstXI linkers"

```

## ORIGIN

```

Query Match      35.8%; Score 521.6; DB 9; Length 842;
Best Local Similarity 77.8%; Pred. No. 1.6e-149;
Matches 655; Conservative 0; Mismatches 184; Indels 3; Gaps 2;

```

```

QY 469 GCACCGGCTAACTACCTGCGCAGACCGCGGTATATCGTAGGGTGCAGACCTTAATCGGA 528
DB 1 GCATCGGCTAACTCTGTGCGCAGACCGCGGTATATCAAGAGATGCAACGTTATCCGGA 60
QY 529 ATTACTGGGCGGTAAAGGTCGCCAGCGGCTTTGTAGTCAGATGGAATCCCGGGCT 588
DB 61 ATGATTGGGCGGTAAAGCGTCTGTAGGTCGTTTAAAGTCCCGTCAATCCAGGGCT 120
QY 589 TAACCTGGGAATTCCTTTGAATCAAGAGCTAGATGTCAGAGGAGGTGAATTC 648
DB 121 CAACCTGGACAGCGCGGTGAAATCAAGCTTAGTACGTAGGGGAGAGGAATTT 180
QY 649 CATGTGTAGCACTGAATATCGTAGAGATGGAAGAACTGATGCGAAGGCAAGCTTC 708
DB 181 CCGGTGAGCGGGTGAATGCGTAGAGATGCGAAGAAACCAACGCGCAAGCACTCTGC 240
QY 709 TGGGTTAACTGACCTCATGACGAAAGCGTGGGAGCAACAGATTAGTACCTTC 768
DB 241 TGGGCGGACACTGACACTGAGAGACGAAGCTAGGAGGAATGAGATTACCCCA 300
QY 769 GTAGTCCAGCGCCCTTAACAGATGCTCACTAGTGTGGGCTTTATAGGCTTG--GTAGC 826
DB 301 GTAGTCTAGCCCTTAACAGATGATAGTACGCTGCTGCTTATGACCCCTGACAGTCTG 360
QY 827 AAGCTTAACGCGGTAAAGTTGACCGCTGGGAGTAGCGGTGCAAGATTAAATCTCAAGGA 886
DB 361 TAGCTTAACGCGGTAAAGTATCCCGCTGGGAGTAGTCCGCAAGAAATGAAATCTCAAGGA 420
QY 887 ATTGACGGGAGACCGCGCAAGCGGTGATATGTGATTAATTCGANTGCAACCGGAAAA 946
DB 421 ATTGACGGGAGACCGCGCAAGCGGTGATGATGTTAATTCGATGCAAGGAAAGAA 480
QY 947 CCTTACCTACCTTTGACATGTAGGCAATTTTCTAGAGATAGATTAGTG--CTTGGGAAACG 1005
DB 481 CCTTACCGAGGCTTTGACATGCGCGGAATCTCTTGAAGAGAGGGGCTCTTGGGAAACG 540
QY 1006 CTAAACAGAGTGTGACATGCGCTGTCTGACGCTGTGAGAGATGTTGGGTTAGTCC 1065
DB 541 CGAGACAGAGTGTGACATGCGCTGTCTGACGCTGTGAGAGATGTTGGGTTAGTCC 600
QY 1066 GCAACGAGCGCAACCTTTGATTAATTTGCAATTTGTTGGGCACTTTAATGAACT 1125
DB 601 GCAACGAGCGCAACCTCTGTGTTAGTTGCAACGTTGAGTTGGAACCTCTGAACAAC 660
QY 1126 GCGGTGACAAACCGGAGGAAGTGGGATGAGTCAAGTCTCATGAGCCCTTAATGGGTA 1185
DB 661 GCGGTGATTAAGCGGAGGAAGTGGGATGAGTCAAGTCTCATGAGCCCTTAATGGCCT 720
QY 1186 GGGCTTCAACATTAATTAATGAGCGCTGACAGAGGTTGCAACCGCGAGAGGGAGCTA 1245
DB 721 GGGGAGACACATGCTCAATGAGCGGAGCAAGAGGTCGAGATCCCGAGAGGGTGAAGTA 780
QY 1246 ATCTCAGAAAGCGGCTGTAGTCCGAGTCCGAGTCTGCAACTGCACTCCGTAAGTGGGA 1305
DB 781 ACTCAGAAAGCGGCTGTAGTCCGAGTCCGAGTCCGAGTCCGTAAGTGGGAAGCGGA 840
QY 1306 AT 1307
DB 841 AT 842

```

```

RESULT 44
BZ483676 837 bp DNA linear GSS 13-DEC-2002
LOCUS BZ483676
DEFINITION BOOAS16T B0.1.6.2_KB_tot Brassica oleracea genomic clone BOOAS16,
genomic survey sequence.
ACCESSION BZ483676
VERSION BZ483676.1 GI:26786074
KEYWORDS GSS
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 837)
Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
15805490
Other_GSSs: BOOAS16T
COMMENT
TIGR
Contact: Chris Town
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: FR
Class: sheared ends.

```

```

FEATURES
Source
Location/Qualifiers
1..837
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOOAS16"
/clone_1ib="B0.1.6.2_KB_tot"
/notes="Vector: pHOSt; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHOSt using BstXI linkers"

```

## ORIGIN

```

Query Match      35.8%; Score 521.4; DB 9; Length 837;
Best Local Similarity 78.0%; Pred. No. 1.9e-149;
Matches 653; Conservative 0; Mismatches 181; Indels 3; Gaps 2;

```

```

QY 469 GCACCGGCTAACTACCTGCGCAGACCGCGGTATATCGTAGGGTGCAGACCTTAATCGGA 528
DB 1 GCATCGGCTAACTCTGTGCGCAGACCGCGGTATATCAAGAGATGCAACGTTATCCGGA 60
QY 529 ATTACTGGGCGGTAAAGGTCGCCAGCGGCTTTGTAGTCAGATGGAATCCCGGGCT 588
DB 61 ATGATTGGGCGGTAAAGCGTCTGTAGGTCGTTTAAAGTCCCGTCAATCCAGGGCT 120
QY 589 TAACCTGGGAATTCCTTTGAATCAAGAGCTAGATGTCAGAGGAGGTGAATTC 648
DB 121 CAACCTGGACAGCGCGGTGAAATCTCAAGCTTGAATCGTAGGGGAGAGGAATTT 180
QY 121 CAACCTGGACAGCGCGGTGAAATCTCAAGCTTGAATCGTAGGGGAGAGGAATTT 180
DB 649 CATGTGTAGCACTGAATGCGTGAAGATATGGAAGAACTGATGCGAAGGCAAGCTTC 708
DB 181 CCGGTGAGCGGTTGAATGCGTGAAGATGCGAAGAAACCAACCGGGAAGCACTCTGC 240
QY 709 TGGGTTAACTGACCTCATGACGTAAGCAAGCGTGGGAGCAACAGATTAGTACCTTC 768
DB 241 TGGGCGGACACTGACACTGAGAGACGAAGCTAGGGGAGCGAATGGAATTATCCCA 300
QY 769 GTAGTCCAGCGCCCTTAACAGATGCTCACTAGTGTGGGCTTTATAGGCTTG--GTAGC 826
DB 301 GTAGTCTAGCCCTTAACAGATGATAGTACGCTGCTGCTATGACCCCTGACAGTCTG 360
QY 827 AAGCTTAACGCGGTAAAGTTGACCGCTGGGAGTAGTCCGCAAGATTAAATCTCAAGGA 886
DB 361 TAGCTTAACGCGGTAAAGTATCCCGCTGGGAGTAGTCCGCAAGAAATGAAATCTCAAGGA 420

```

QY 887 ATTGACGGGGACCCGACAGAGGGGTGATGATTTAATTCATGCAACGGGAAA 946  
Db 421 ATTGACGGGGACCCGACAGAGGGGTGATGATTTAATTCATGCAACGGGAAA 480  
QY 947 CCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 1005  
Db 481 CCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 540  
QY 1006 CTAAACAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065  
Db 541 CGGACACAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
QY 1066 GGAAGAGGACCACTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1125  
Db 601 GGAAGAGGACCACTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 660  
QY 1126 GCGGATGACCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185  
Db 661 GCGGATGACCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
QY 1186 GGGCTTACACGTAATACATGATGATGATGATGATGATGATGATGATGATGATGAT 1245  
Db 721 GGGCTTACACGTAATACATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
QY 1246 ATCTCAGAAAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1302  
Db 781 ATCTCAGAAAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 837

RESULT 45  
CP557211  
LOCUS 770 bp mRNA linear EST 23-SEP-2003  
DEFINITION 1115025D01.x1 C. reinhardtii CC-1690 (mt+), Stress III  
(normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 3', mRNA  
sequence.  
ACCESSION CP557211  
VERSION CP557211.1 GI:34983294  
KEYWORDS EST  
SOURCE Chlamydomonas reinhardtii  
ORGANISM Chlamydomonas reinhardtii  
Chlamydomonas reinhardtii  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;  
Chlamydomonadales; Chlamydomonadales; Chlamydomonas.  
REFERENCE 1 (bases 1 to 770)  
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C.,  
Merchant, S., Quinn, J., and Shrager, J.  
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
Unicellular System for Analyzing Gene Function and Regulation in  
Vascular Plants. Project: 1115  
JOURNAL Unpublished (2003)  
COMMENT Contact: Charles Hauser  
DCMB Box 91000  
Duke University  
Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: chausser@duke.edu  
High quality sequence ecop: 770.  
FEATURES  
source  
1..770  
/organism="Chlamydomonas reinhardtii"  
/mol\_type="mRNA"  
/strain="CC-1690 wild type mt+ 21gr"  
/db\_xref="taxon:3055"  
/clone\_1ib="C. reinhardtii CC-1690 (mt+), Stress III  
(normalized), Lambda Zap II"  
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:  
XhoI; Stress condition III library, constructed by  
Jeanette Quinn and Chung-Men Chang, combines cDNAs from  
CC-1690 cells grown to mid-log phase in copper-free YAP  
medium (see Quinn and Merchant (1998) Methods in  
Enzymology, 2997:263-279) in a shaking (250 rpm) illuminated  
(about 100 micromole/m2/sec) incubator at 22 C (see Quinn

and Merchant (1998) Methods in Enzymology, 2997:263-279);  
CC-1690 cells grown to mid-log phase in low Fe (1  
micromolar Fe) YAP medium (see La Fontaine S, Quinn JM,  
Nakamoto S, Page MD, Gohre V, Moseley JJ, Kropac J,  
Merchant S, LaFontaine et al. (2002) Eukaryotic Cell,  
1:736-757) in a shaking illuminated incubator (same  
conditions as above). CC-1690 cells were grown to mid-log  
phase in YAP medium in a shaking illuminated incubator to  
a density of 8x10<sup>6</sup> cells/ml. The flask was transferred to  
a shaking platform (200 rpm) at room temperature (23C) 12  
micromole/m2/sec illumination and bubbled in a stoppered  
flask with 98% nitrogen, 2% CO<sub>2</sub> gas mixture for 1 hour  
prior to harvesting for RNA isolation (see per Quinn JM,  
Barraco P, Eriksson M, Merchant S, Quinn et al. (2000) JBC  
275:6080-6089); CC-1690 cells grown to mid-log phase  
(3x10<sup>6</sup> cells/ml) in YAP medium in a shaking (150 rpm)  
illuminated (70 mole photon/m2/sec) incubator at 27 C.  
Cells were diluted to 1x10<sup>6</sup> cells/ml, transferred to high  
light (11000 mol photon/m2/s) with shaking (150 rpm) and  
sampled at (0.5, 1.2, 4.6, 12 hrs). PolyA mRNA was purified  
from each sample, pooled and cDNA synthesized (see Shrager  
et al., Plant Physiol. 131, 401-408 for details). The cDNA  
was directionally cloned into Lambda Zap II (Stratagene)  
in the EcoRI (5') and XhoI (3') sites. pBluescript II SK-  
plasmids were excised from the lambda Zap clones by  
superinfection with ExAssist (Stratagene) phage. The  
library was normalized using method 4 described in Bonaldo  
et al., (1996) Genome Research 6: 791-806."

ORIGIN  
Query Match 35.7%; Score 520.6; DB 6; Length 770;  
Best Local Similarity 84.3%; Pred. No. 3.3e-149;  
Matches 633; Conservative 0; Mismatches 114; Indels 4; Gaps 4;

QY 468 AGCACCGGCTTAACCTACCTGACGACGACCGCGGTAACTAGTGGTGCAGAGCTTAATCGG 527  
Db 21 AACACCGGCTTAACCTGACGACGACCGCGGTAACTAGTGGTGCAGAGCTTAATCGG 80  
QY 528 AATTACTGGGGGTAAGGGGTGCGAGCGGCTTTGTAAGTGAATGTAATCCCGGGC 587  
Db 81 AATTACTGGGGGTAAGGGGTGCGAGCGGCTTTGTAAGTGAATGTAATCCCGGGC 140  
QY 588 TTAACCTGGGAATTCGTTTGAACCTACAGCTAGAGTGTGGCAGAGAGGTGAATT 647  
Db 141 TCAACTGGGAATTCGTTTGAACCTAGTGTGGCAGAGAGGTGAATT 200  
QY 648 CCATGTGTAGAGTAAGTAATGCTAGATATGAAAGAACATCATGATGCGGAAGGCACTTC 707  
Db 201 CCAGGTGTAGAGTAAGTAATGCTAGATATGAAAGTAATACCGGTGGGAAGGCGGCC 260  
QY 708 CTGGGTTAACTACGCTACGCTACGACGAAAGCGTGGGAGCAACAGGATTAGATACCT 767  
Db 261 CTGGGTTAACTACGCTACGCTACGACGAAAGCGTGGGAGCAACAGGATTAGATACCT 320  
QY 768 GGTAGTCCACGCGCTAAACGATGTCACTAGTGTGGGCTTAAAGCTTG-7AACG 826  
Db 321 GGTAGTCCACGCGCTAAACGATGTCACTAGTGTGGGCTTAAAGCTTG-7AACG 380  
QY 827 AAGCTAACGCTGAAAGTGAACCGCTGGGAGTACCGTGCAGAAATTAAACTCAAGAGA 886  
Db 381 GAGCTAACGCTTAAAGTGAACCGCTGGGAGTACCGTGCAGAAATTAAACTCAAGAGA 440  
QY 887 ATTGACGGGGACCCGACAGAGGGGTGATGATTTAATTCATGCAACGGGAAA 946  
Db 441 ATTGACGGGGACCCGACAGAGGGGTGATGATTTAATTCATGCAACGGGAAA 500  
QY 947 CCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 1005

Db 501 CCTTACCTGCTTGCATCCACGGAAGTTTCAAGATGAGATGTGCTTCGGGAACC 560  
Qy 1006 CTACACAGGTGCTGCATGAGTGTGCTGAGTGTGCTGAGATGTTGGTTAAGTCCC 1065  
Db 561 GTGACACAGGTGCTGCATGAGTGTGCTGAGTGTGCTGAGATGTTGGTTAAGTCCC 620  
Qy 1066 GCACAGAGCGCAACCTGTGCTAATTAATGCTATC-ATTGTTGGGCACTTAAATGACAC 1124  
Db 621 GCACAGAGCGCAACCTGTGCTAATTAATGCTATC-ATTGTTGGGCACTTAAATGACAC 680  
Qy 1125 TCCCGGTGACAAACCGGAGGAGTGGGAGTACGTCAGTCCCTAATGAGGCT 1184  
Db 681 TCCCGGTGACAAACCGGAGGAGTGGGAGTACGTCAGTCCCTAATGAGGCT 739  
Qy 1185 AGGCGTTCACAGTAAATACATGCGCGCTAC 1215  
Db 740 AGGCGTTCACAGTAAATACATGCGCGCTAC 770

RESULT 46  
BH545955/c 866 bp DNA linear GSS 14-DEC-2001  
LOCUS BOGIX59TF BOGI Brassica oleracea genomic clone BOGIX59, genomic  
DEFINITION survey sequence.

ACCESSION BH545955  
VERSION BH545955.1 GI:17817795  
KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 866)  
Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Utterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.

Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)

JOURNAL PUBMED 15805490  
COMMENT Other\_GSSs: BOGIX59TR  
Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208

Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.

FEATURES  
Location/Qualifiers  
1..866

/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO100DH3"  
/db\_xref="taxon:3712"  
/clone="BOGIX59"  
/clone\_1lb="BOGI"  
/note="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared  
genomic DNA inserted into pHOSt using BstXI linkers"

## ORIGIN

Query Match 35.6%; Score 519.2; DB 9; Length 866;  
Best Local Similarity 77.6%; Pred. No. 9,2e-149;  
Matches 666; Conservative 0; Mismatches 188; Indels 4; Gaps 3;  
Qy 584 GGGCTTAACCTGGGAATGCGTTGAAACTACAGAGCTAGAGTGTGCAGAGGAGGCTGG 643  
Db 866 GGGCTTAACCTGGGAATGCGTTGAAACTACAGAGCTAGAGTGTGCAGAGGAGGCTGG 807  
Qy 644 AATTCATGTGTAGCACTGAAATCGCTAGAGATGTGAAGAAACATTCAGTGGCGAAGGAG 703  
Db 806 AATTCATGTGTAGCACTGAAATCGCTAGAGATGTGAAGAAACATTCAGTGGCGAAGGAG 747

Qy 704 CTTCTGGGTTAAACATGACGCTCATGACGAAAGCGTGGGAGACAAACAGATTAGATA 763  
Db 746 TCTGCTGGGCGGACATCTTACATGAGAGACGAAAGCTAGGGAGCGAATGGGATTAGATA 687  
Qy 764 CCTGTGATGCCACCGCTTAAACGATGTCACTAGTGTGGGCTTATTAAGCTTG--G 821  
Db 666 CCCGATGTGCTTACCGCTTAAACGATGTCACTAGTGTGGGCTTATTAAGCTTG--G 627  
Qy 822 TAACGAAGCTAACCGGTGAATTTGACCGGCTGGGAGTACGGTGGCAAGATTAAACTCA 881  
Db 626 TCTCTAGCTTAACCGGTGAATTTGACCGGCTGGGAGTACGGTGGCAAGATTAAACTCA 567  
Qy 882 AAGGAATTGACGGGAGACCGGACCAAGCGGTGATTAATGATTAATGATGACACGG 941  
Db 566 AAGGAATTGACGGGAGACCGGACCAAGCGGTGATTAATGATTAATGATGACAAAGG 507  
Qy 942 AAAAAGCTTAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 1000  
Db 506 AAGAACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 447  
Qy 1001 GAACGCTAACACAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060  
Db 446 GAACGCTAACACAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387  
Qy 1061 GTCCCGCAACGAGCGGCAACCTTGTATTAATGATTAATGATTAATGATTAATGATTAATG 1120  
Db 386 GTCCCGCAACGAGCGGCAACCTTGTATTAATGATTAATGATTAATGATTAATGATTAATG 327  
Qy 1121 AGACTGCCGCTGACAAACCGGAGGAGTGGGAGTACGTCAGTCCCTGATGAGCTTAT 1180  
Db 326 AAGTCTCCGCTGATTAAGCGGAGGAGTGGGAGTACGTCAGTCCCTGATGAGCTTAT 267  
Qy 1181 GGGTGAAGGCTTACACAGTAAATCAATGCGGCTGACAGAGGTTGCCAACCCCGAGAGGG 1240  
Db 266 GCCCTGGGCGACACAGTGTCTCAATGCGGAGCAAAAGTGTGCGATCCCGAGAGGGT 207  
Qy 1241 AGTAACTTCAAGAAAGCGGCTGATGCTGCGAGTTCGGAATGCACTGCACTGCGTGAAG 1300  
Db 206 AGTAACTTCAAGAAAGCGGCTGATGCTGCGAGTTCGGAATGCACTGCACTGCGTGAAG 147  
Qy 1301 TCGGAATCGCTAGTAAATCGCGGATCAG-CATGTCGCGGTGAATACGTTCCCGGCTTGT 1359  
Db 146 CCGGAATCGCTAGTAAATCGCGGATCAG-CATGTCGCGGTGAATACGTTCCCGGCTTGT 87  
Qy 1360 ACAACACCGCGCTGACACCATGAGAGTGGTTCACCAAGACGATGCTTAAACCTTAA 1419  
Db 86 ACACACCGCGCGTCACTATGAGAGCTGGCCATGCGCAAGTCTTAACTTAAACCTTAA 27  
Qy 1420 GAGAGGCGCTTGGCAGCG 1437  
Db 26 GAGAGGCGCTTGGCAGCG 9

RESULT 47  
BH545929 849 bp DNA linear GSS 14-DEC-2001  
LOCUS BOGNI60TF BOGN Brassica oleracea genomic clone BOGNI60, genomic  
DEFINITION survey sequence.  
ACCESSION BH545929  
VERSION BH545929.1 GI:17797710  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 849)  
Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Utterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
JOURNAL PUBMED 15805490

## COMMENT

Other GSSs: BOMND50TR  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdclown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TP  
 Class: sheared ends.

## FEATURES

source

Location/Qualifiers

1..849  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="T01000DH3"  
 /db\_xref="taxon:3712"  
 /clone\_1lb="BOMND50"  
 /clone\_1lb="BOMND50"  
 /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHOS1 using BstXI linkers"

## ORIGIN

Query Match 35.6%; Score 518.8; DB 9; Length 849;  
 Best Local Similarity 78.1%; Pred. No. 1.2e-148;  
 Matches 662; Conservative 0; Mismatches 182; Indels 4; Gaps 3;

QY 516 AGCGTTAATCGGAATTACTGGCGCTGAAGAGGTCGCGAGCGGCTTTGTAAGTCAGATGTG 575  
 DB 2 AGCGTTATCCGGAATGATGGCGCTGAAGCGCTGTAGGTGCGCTTTTAAGTCCGCCGTC 61  
 QY 576 AAATCCCGGGGCTTAACCTGGGAATTTGGCTTGAACCTACAGAGGCTAGAGTGGCAAG 635  
 DB 62 AAATCCCAAGGGCTCAACCTGGACAGGCGCTGGAACCTACAGGCTTGAAGTACGTAAGG 121  
 QY 636 GAGGTGGAATTCATGTGTACAGTGAATCGTAGAGATATGAAGAATCATCATGATGTC 695  
 DB 122 GAGAGGGAATTTCCGGTGAAGCGGTGAATGCTAGAGATCGAAGAATCAACAGCGC 181  
 QY 666 GAGGCAAGCTCTCTGGGTTAACTGACCTCATGACGAAAGCGTGGGAGCAAAACAGG 755  
 DB 182 GAAACACACTCTGCTGGCCGACACTGACACTGACGAAAGCTAGGGAGCGAATGGG 241  
 QY 766 ATTAGATCCCTGTAGTCCAGCCCTTAAGCATGTAAGTGAATGTTGGGCTTATTTAG 815  
 DB 242 ATTAGATCCCTGTAGTCCAGCCCTTAAGCATGTAAGTGAATGTTGGGCTTATTTAG 301  
 QY 816 GCTTG--GTACGAAGCTAACCGCTGAAGTTGACCGCTGGGAGATCGGTGCAAGATT 873  
 DB 302 CCGTCAAGTGTGTAGCTAACCGCTGAAGTTGACCGCTGGGAGATCGGTGCAAGATT 361  
 QY 874 AAAAATCAAGGAATTGACGGGAGACCCGACAAAGCGGTGATTAATGTTGATTAATTCAT 933  
 DB 362 GAAACTCAAGGAATTGACGGGAGACCCGACAAAGCGGTGATTAATGTTGATTAATTCAT 421  
 QY 934 GCAAGCGCAAAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 993  
 DB 422 GCAAGCGCAAAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 481  
 QY 994 G-CTTCGGGAACGCTAACACAGGTGCTGATGCTGTCTGACCTGCTGTCTGATGATGT 1052  
 DB 482 GCTTCGGGAACGCTAACACAGGTGCTGATGCTGTCTGACCTGCTGTCTGATGATGT 541  
 QY 1053 TGGGTTAAGTCCGCAACGAGCGCAACCTTGTCTAATTAATTCATCATTTGGTGGCA 1112  
 DB 542 TGGGTTAAGTCCGCAACGAGCGCAACCTTGTCTAATTAATTCATCATTTGGTGGCA 601  
 QY 1113 CTTTATATGAGTCCCGTGGAACACCGGAGGAAGTGGGATGACCTCAATCTCTCATG 1172  
 DB 602 CCTTAACAGACTGCGCGGTATAGCCGAGGAAGGTGAGATGACGTAAATGTCATG 661  
 QY 1173 GCGCTTATGAGGCTTCAACAGTATATACATGGCGGTGACAGAGGTTGCCAACCG 1232  
 DB 662 CCCCTTATGCGCTGGGCGACACGTGTCTAATGCGCGGAGCAAAAGGTTCCGATCCCG 721

QY 1233 CGAGGGGAGCTAATCTCAGAAAGCGCGTCTGAAGTCCGATCGAGTCTGCACTGCACT 1292  
 DB 722 CGAGGTAGCTAATCTCAGAAACCCGCTCTCAAGTTCGATTCAGAGCTGCACTCGCT 781  
 QY 1293 CCGTGAAGTCCGAATTCGCTAGTAAATCGCGATCAG-CATGTCGCGGTGAATACGTTCCG 1351  
 DB 782 GCATGAAGCCCGGATTCGCTAGTAAATCGCGGTGAGCCATACGCGGTGAATTCGTTCCG 841  
 QY 1352 GGTCTTGT 1359  
 DB 842 GGCCTTGT 849

RESULT 48  
 BH700021/c 995 bp DNA linear GSS 20-FEB-2002  
 LOCUS BOMND50TR BO.2.3 KB Brassica oleracea genomic clone BOMND50,  
 DEFINITION genomic survey sequence.  
 ACCSSION BH700021  
 VERSION BH700021.1 GI:18775982  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 995)  
 Ayle, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
 Utecher, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 15805490  
 Other GSSs: BOMND50TR  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdclown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.

FEATURES  
 source  
 1..995  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="T01000DH3"  
 /db\_xref="taxon:3712"  
 /clone\_1lb="BOMND50"  
 /clone\_1lb="BO.2.3 KB"  
 /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Query Match 35.6%; Score 518.4; DB 9; Length 995;  
 Best Local Similarity 74.7%; Pred. No. 1.7e-148;  
 Matches 731; Conservative 0; Mismatches 208; Indels 39; Gaps 5;

QY 64 GTGGGAGTGGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 123  
 DB 942 GTTTCAGTGGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 883  
 QY 124 CATCAAAAGATGTGTTAATACCGCATATATCTTAAGAGAAAGCAGGAGATCGAAAGAC 183  
 DB 882 GGTGAAACCGGTGTGTTAATACCGCATATATCTTAAGAGAAAGCAGGAGATCGAAAGAC 832  
 QY 184 CTGGGCTTTTGGAGCGGCGCATGTCTGATTAAGTATGTTGGTGAAGGCTTAAAGCTTACCA 243  
 DB 831 ----CGCCCGAGAGGGGCTCGCGCTGATTAAGTATGTTGGTGAAGGCTTAAAGCTTACCA 776  
 QY 244 GCGCAGATCAATGTTGTGTGAGAGAGACGACGACACTGGGACTGAGACAGCGCC 303

```

Db      ||||| 775 GGCAGTGAATCAGTACGTGTGCTCCAGAGGATGATCACCACACTGGGACTGAGACACGGCC 716
Qy      304 CAGACTCTACGGGAGCAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCCTGATCCAG 363
Db      715 CAGACTCTCTACGGGAGCAGCAGTGGGGAATTTTCCGCAATGGGCGAAGCCTGATCCAG 656
Qy      364 CAATGCCCGGTGAGTGAAGAGGCGCTTGGGTTTGAAGCTCTTTCAGTGCAGAGAGAAA 423
Db      655 CAATGCCCGGTGAGTGAAGAGGCGCTTGGGTTTGAAGCTCTTTCAGTGCAGAGAGAAA 599
Qy      424 GGTACCGTAAATATATCTGACTCATGACGGTATCCAGAGAGAGAACACCGGTATCTAC 483
Db      598 -----AGCAATGACGGTATCTGGGGAATTAACATCGGTAACTCT 559
Qy      484 GTGCCAGACGCCCGGTAAATACGTAGAGGTGCAAGCCTTAATCGGAATTAATCTGGGCGTAA 543
Db      558 GTGCCAGACGCCCGGTAAATACAGAGATGCAAGCCTTAATCGGAATTAATGGGCGTAA 499
Qy      544 GGGTGGCAGGCGCGCTTTTGAAGTCAAGTGTGAATCCCGGGCTTAACTGGGAAATTGC 603
Db      498 GCGTCTGTAGTGGCTTTTAAAGTCCGCGCTCAAAATCCAGGCGTCAACCTGAGACAGG 439
Qy      604 GTTGAATACTCAAGGCTAGAGTGTGSCAGAGGAGGTGAATTCATGTGTAGCAGTGA 663
Db      438 GGTGAAACTACCAAGCTTGAATGATGAGGCGAGGGAATTTCCGGTGAAGCGGTGA 379
Qy      664 AATGCCGTAGATATGAGGAACATCGATGAGCGAGCGAGCCTCTGGGTTTAACTGATGAC 723
Db      378 AATGCCGTAGATATGAGGAACATCAACCGGCAAGGACCTCTGCTGGCGGACACTGAC 319
Qy      724 GCTCATGACGAAAGCGTGGGAGCAAAACAGATTGATACCTGTGATGTCACGCGCTTA 783
Db      318 ACTGAGAGAGAAAGTGAAGGAGGAGCGAATGGATTGATACCCAGATGCTTAAAGCGTGA 259
Qy      784 AACGATGTCAACTGATTTTGGGCGCTTATTAAGCGTTG--GTAAAGAACTAAAGCGGTGA 841
Db      258 AACGATGTCAACTGATTTTGGGCGCTTATTAAGCGTTG--GTAAAGAACTAAAGCGGTGA 199
Qy      842 GTTGACCGCCTGGGAGTACGGTCCGAAGATTAAATCAAGAGAAATGAAGGAGACCG 901
Db      198 GTATCCCGCTGGGAGTACGGTCCGAAGATTAAATCAAGAGAAATGAAGGAGACCG 139
Qy      902 CACAAGCGGTGATTAATGATTAATTCATGATCAAGCAAGCAAAACCTTAACTCACTCTG 961
Db      138 CACAAGCGGTGATTAATGATTAATTCATGATCAAGCAAGCAAAACCTTCAAGGCTTG 79
Qy      962 ACATGTAGCGAATTTTCTAGAGATGATTAAGTG-CTTGGGAAACGTTAAACAGGTGCTG 1020
Db      78 ACATGCCCGCAATCTCTTGAAGAAGAGAGGGTGCCTTCGAAACGCGAGACAGGTGCTG 19
Qy      1021 CATGGCTGTCTCAGCTC 1038
Db      18 CATGGCTGTCTCAGCTC 1

```

```

RESULT 49
BH685417 1020 bp DNA linear GSS 19-FEB-2002
LOCUS BH685417 BOMFG90TR BO_2_3_KB Brassica oleracea genomic clone BOMFG90,
DEFINITION genomic survey sequence.
ACCESSION BH685417
VERSION BH685417.1 GI:18755854
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; euroside II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 1020)
Aylele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utecherack,T.R., Wortman,J.R., White,O.R. and Town,C.D.
Whole genome shotgun sequencing of Brassica oleracea and its

```

JOURNAL  
PUBMED  
15805490  
COMMENT  
Other\_GSSs: BOMFG90TR  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source  
1..1020  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone\_1lb="BO\_2\_3\_KB"  
/note="vector: pHD51; Site\_1: BactX; 2-3 kb sheared  
genomic DNA inserted into pHD51 using BactX linkers"

## ORIGIN

Query Match 35.5%; Score 517.6; DB 9; Length 1020;  
Best Local Similarity 73.8%; Pred. No. 3e-148;  
Matches 742; Conservative 0; Mismatches 224; Indels 40; Gaps 5;

```

Qy      64 GTGGCAGAGTGGGCGAGCGGCTGAGTAAATGATCGAAGCTATCCAGAGAGGGGGGTAAAG 123
Db      47 GTTTCAGATGGCGAGCGGCTGAGTAAATGATCGAAGCTATCCAGAGAGGGGGGTAAAG 106
Qy      124 CATCGAAGATGTCTAATACCGCATATATCTTAAGAGAGAAACAGGGGATGAAAGAC 183
Db      107 GGTGAAACGGGCTGCTAATACCCGTAGG--CTGAGAGACAAAGAGGAATC----- 157
Qy      184 CTTCGCGCTTTTGAAGCGCGCATGTCTGATTTAGCTTGTGTGGGTTAAAGCTTACCA 243
Db      158 ----CCCCGAGAGAGGGGCTCGCTGTAGTATGCTTGTGTAGGCAATAGCTTACCA 213
Qy      244 GCGCAGATCAGTATGTGCTGAGAGAGACCAAGCCTGAGTGAACAGCGGC 303
Db      214 GCGCAGATCAGTATGTGCTGAGAGAGATGATCAGCCACACTGGGACTGAGACAGCGGC 273
Qy      304 CAGACTCTTACCGGAGCAGCAGATGGGGAATTTTGGACAATGGGCGCAAGCCTGATCCAG 363
Db      274 CAGACTCTTACCGGAGCAGCAGATGGGGAATTTTCCGCAATGGGCGAAGCCTGAGCAG 333
Qy      364 CAATGCCCGGTGAGTGAAGAGGCGCTTGGGTTTGAAGCTCTTTCAGTGCAGAGAGAAA 423
Db      334 CAATGCCCGGTGAGTGAAGAGGCGCTTGGGTTTGAAGCTCTTTCAGTGCAGAGAGAAA 390
Qy      424 GGTACCGTAAATATATCTGACTCATGACGGTATCCAGAGAGAGAACACCGGTATCTAC 483
Db      391 -----AGCAATGACGGTATCTGGGGAATTAACATCGGTAACTCT 430
Qy      484 GTGCCAGACGCCCGGTAAATACGTAGAGGTGCAAGCCTTAATCGGAATTAATCTGGGCGTAA 543
Db      431 GTGCCAGACGCCCGGTAAATACAGAGATGCAAGCCTTAATCGGAATTAATGGGCGTAA 490
Qy      544 GGGTGGCAGGCGCGCTTTTGAAGTCAAGTGTGAATCCCGGGCTTAACTGGGAAATTGC 603
Db      491 GCGTCTGTAGTGGCTTTTAAAGTCCGCGCTCAAAATCCAGGCGTCAACCTGAGACAGG 550
Qy      604 GTTGAATACTCAAGGCTAGAGTGTGSCAGAGGAGGTGAATTCATGTGTAGCAGTGA 663
Db      551 GGTGAAACTACCAAGCTTGAATGATGAGGCGAGGGAATTTTCCGGTGAAGCGGTGA 610
Qy      664 AATGCCGTAGATATGAGGAACATCGATGAGCGAGCGACCTCTGTGGTTAACTGATGAC 723
Db      611 AATGCCGTAGATATGAGGAACATCAACCGGCAAGGACCTCTGTGGGCGGACACTGAC 670

```



QY	724	GCTCATGCACGAAACACGTCGGGAGCAAAACAGGATTAGATACCCCTGTATGTGCCAGCCCTA	783		
QY	724	GCTCATGCACGAAACACGTCGGGAGCAAAACAGGATTAGATACCCCTGTATGTGCCAGCCCTA	783		
DB	671	ACTGAGAGACGAAACCTAGGGGAGGAATGGGATTAGATACCCAGTATGTCTTACCCGTA	730		
QY	784	AACGATGTCAATTAAGTATGTTGGGCGCTTATTAGGCTTGG--GTACAGAAAGCTTAACGCGTGAA	841		
DB	731	AACGATGTCAATTAAGTATGTTGGGCGCTTATTAGGCTTGG--GTACAGAAAGCTTAAACGCGTTAA	790		
QY	842	GTTGACCGCTGGGGAGTAGTCGGTCGCAAGATTAAACTCAAAAGAAATTGACGGGGACCCG	901		
DB	791	GTATCCCGCTGGGAGTAGTCGTTCCCAAGAAATGAAGCTCAAGAAATTGACGGGGACCCG	850		
QY	902	CACAAACCGGTGGATTATGTGGATTATTTGATGATGCAACGGGAAAACTTAACTTACCTACCTTTG	961		
DB	851	CACAAACCGGTGGAGCATGTGTGTTAAATTCGAATGCAAAAGGAAACCTTAAACAGGCGCTTG	910		
QY	962	ACATAGTAGCGAAT--TTTCTAGAGATAGATTTAGTGTGCTTCGGGAACGCTAACACAGGTCT	1011		
DB	911	ACATGCCCGCAATCCCCCTTTGAAGAGAGGGTGCCTTTGGAAACGGCAACACAGTGTGT	970		
QY	1030	GCATGCGCTGTCTGTCAGCTCTGTCTGTGAGATGTTGGGTTAAAGTCCC	1065		
DB	971	GCATGCGCTGTCTGTCAGCTCTGTCTGTGAGATGTTGGGTTAAAGTCC	1016		
RESULT 50	BH730827	848 bp	DNA	linear	GSS 20-FEB-2002
LOCUS	BOMGL56TF	BO_2_3	KB	Brassica oleracea	genomic clone BOMGL56,
DEFINITION	genomic survey sequence.				
ACCESSION	BH730827				
VERSION	BH730827.1	GI:19836222			
KEYWORDS	GSS.				
SOURCE	Brassica oleracea				
ORGANISM	Brassica oleracea				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.				
AUTHORS	1 (bases 1 to 848)				
TITLE	Ayele, M., Haas, B. J., Kumar, N., Mu, H., Xiao, Y., Van Aken, S., Utecherback, T. R., Wortman, J. R., White, O. R., and Town, C. D.				
JOURNAL	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis				
PUBMED	Genome Res. 15 (4), 487-495 (2005)				
COMMENT	15805490				
CONTACT	Chris Town				
TELEPHONE	9712 Medical Center Drive, Rockville, MD 20850, USA.				
TELEFAX	Tel: 301-838-3523				
TELEFAX	Fax: 301-838-0208				
EMAIL	Email: cdtown@tigr.org				
NOTE	DNA is from a doubled haploid provided by Tom Osborn.				
SEQ PRIMER	Seq primer: TP				
CLASS	Class: sheared ends.				
LOCATION	Location/Qualifiers				
FEATURES	1..848				
SOURCE	/organism="Brassica oleracea"				
	/mol_type="genomic DNA"				
	/strain="TO1000DH3"				
	/db_xref="taxon:3712"				
	/clone="BOMGL56"				
	/clone_1lb="BO_2_3_KB"				
	/note="Vector: pHOSt1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt1 using BstXI linkers"				
ORIGIN					
Query Match	35.5%;	Score 517;	DB 9;	Length 848;	
Beet Local Similarity	77.5%;	Pred. No. 4,4e-148;			
Matches 675;	Conservative	0;	Mismatches 170;	Indels 26;	Gaps 3;
QY	208	TCTGATTAGCTAGTGTGGTGGGTAAGGCTTACCAAGCGACGACGATCAGTATGTTGCTGCA	267		
DB	848	TCTGATTAGCTAGTGTGGTGGGTAAGGCTTACCAAGCGATCAGTATGTTGCTGCA	789		

OY		268	GAGGACGCACCAGCCACACTCGGGACCTGAGACAACACGCCCAAGCATCTTCATCCGGAAAGCAGCACGT	327
Db		788	AAGGATGATCAGCCACACTGGGAGTCTGAGACACAGGCCCAAGCATCTTCATCCGGAAAGCAGCACGT	729
OY		328	GGGGAATTTTTGGAACATGGGCGCAAGCCTGATCCAGCAATGCCCCGTGAGTAGTAAGAAGGC	387
Db		728	GGGGAATTTTCCGCAATTGGGGGAAAGCCTGACAGAGCAATGCCCCTGGAGAGGTAGTAAGAAGGC	669
OY		388	CTTCGGGTGTGAAGCTCTTTCAGTCGAGAAAGAAAAGGTTACGGTAAATAATCGTGACTC	447
Db		668	CTACCGGCTCCGAACTCTTTTCCAGAGAAAG-----AGCA	632
OY		448	ATGACGGTATCGACAGAAAGACACGGGCTAATCTACGTGCCAGACCGCGGTAAATACGT	507
Db		631	ATGACGGTATCTGGGGAAATTAAGCATTCGGTACTCTGTGCCAGACCGCGGTAAATACAG	572
OY		508	AAGGTGCAAGCGTTAATCGGAATTACTGGCGCTAAAGGGGTGCGACGCGGCTTTGTAAGT	567
Db		571	AGGATGCAAGCGTTATCCGGAATGATTTGGGCGTAAAGSGTCTGTAGTGCGCTTTTAATGT	512
OY		568	CAGATGTGAATATCCCCGGGCTTAACTCTGGGAATTGGCGTTTGAAATCAAGAGGCTAGAGTG	627
Db		511	CCGCGGTAAATATCCCAAGGCTCTMACCCCTGGAAGGCGGTGGAACCTAACCAAGCTTGAGTA	452
OY		628	TGGCAGAGGGAGGTGGAATTCATGTGTAGCAGTGAATAATGGCTAGATGATATGAAGACA	687
Db		451	CGGTAGGGGCGACAGGGAATTTTCGGTGGAGCGGTGAATATGCTTAGAGATTCGAAAGAAC	392
OY		688	TCGATGGCGAAGGAGCAGCCTCCTGGGTTAAACTGACGCTCATGACGAAAGCTGGGGAG	747
Db		391	CCAACGGCGAAGACACTCTGCTGGGCGCACACTGACCTGAGAGACGAAAGCTTAGGGGGG	332
OY		748	CAAACAGGATTAAGATACCTGTGTAGTCCAGCGCCCTAAACGATGTCACTAGTGTGTGGGC	807
Db		331	CGAATGGGATTAAGATACCCACGTAGTCTTAGCGCTGAACGATGATCTTAGGCGCTGTGTC	272
OY		808	CTTATTAGAGCTTGG--GTAAACGAAGCTAACGCGTGAAAGTTGAACCGGCTGGGAGTACGCTC	865
Db		271	GTATCGACCCGTCGAGCTGCTGTAGCTTAACGGGTTAAGTATCCCGGCTGGGGAGTACGTTTC	212
OY		866	GCAAGATTAAACTCAANAGAAATTGACGGGGACCCGCAACAGCGGTGATTAATGTGATTT	925
Db		211	GCAAGAAATGAAGAACTCAAGAAATTTGACGGGGGCCGCAACAGCGGTGAGCATGTGGTTT	152
OY		926	AATTCGATGCAACGCGGAAAACTTACCTACCTGATGATGAGGAATTTTCTTAGAGAT	985
Db		151	AATTCGATGCAAGGAGAAACCTTACCAAGGGCTTTGACATGCGCGMAATCCTCTTGAAAG	92
OY		986	AGATTAGTG-CTTCGGGAACGCTAACACAGAGTCTGCAATGCTGTCTGTCAGTCTGTGTCG	1044
Db		91	AGAGGGGTCCTTGGGGAACGCGACACAGGGTGGTGCATGGCTGTCTGTCAGTCTGTGCGG	32
OY		1045	TGAGATGTTGGGTTAAGTCCCGCAACGACGC	1075
Db		31	TAAAGTGTTGGGTTAAGTCCCGCAACGACGC	1
<hr/>				
RESULT 51				
BH535640/c				
LOCUS	BH535640	848 bp	DNA	linear GSS 14-DEC-2001
DEFINITION	BOOGL94TF BOGY Brassica oleracea genomic clone BOOGL94, genomic survey sequence.			
VERSION	BH535640			
KEYWORDS	BH535640.1 GI:17766867			
ACCESSION	GSS.			
SOURCE	Brassica oleracea			
ORGANISM	Brassica oleracea			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosid; eurosid II; Brassicales; Brassicaceae; Brassica.			
AUTHORS	Ayalew,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,			



Uterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of *Brassica oleracea* and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490  
PUBMED  
Other\_GSSs: B0GYL94TR  
Contact: Chris Town  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

# FEATURES

Location/Qualifiers  
1..848  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO100DH3"  
/db\_xref="taxon:3712"  
/clone="B0GYL94"  
/clone\_1ib="B0GY"  
/note="Vector: PHOS1; Site\_1: BstXI; 2-3 kb sheared  
genomic DNA inserted into PHOS1 using BstXI linkers"

## ORIGIN

Query Match 35.4%; Score 516.6; DB 9; Length 848;  
Best Local Similarity 77.1%; Pred. No. 5.8e-148;  
Matches 661; Conservative 0; Mismatches 184; Indels 4; Gaps 3;

449 TGACGGTATCGACAGAAAGACCGGCTTACTACCTGCGACAGCGCGGTATATACGTA 508  
848 TACCGTATCTGGGAAATAGCATCGCTTACTCTGTGCCAGACGCGGTATATACAG 789  
509 GGGTCGACGCTTATCGAATTACTGGGCGTAAAGGGTGCAGCGCGCTTTGTAAGTC 568  
788 GGATGCAAGGTTATCCGGATGATTGGGCGTAAACGCTGTAGTGGCGCTTT-TAAGTC 730  
569 AGATGTGAATCCCGGGCTTAACTCGGGAATTCGTTGAACCTCAAGGCTAGAGTGT 628  
729 CGCGCTCAATCCAGGGCTCAACCTTGACAGCGGCTGGAATCTTCAAGCTTGAATAC 670  
629 GGCAGAGGGAGGTAATTCATGTGTAGAGTGAATGCGTGAAGATGAGAAACAT 688  
669 GGTAGGGGCAAGAGGAAATTTCCGCTGAGCGGTAATGTCTAGAGATCGGAAGAACA 610  
689 CGATGCGCAAGGACGCTCTGGGTTAACACTGACGCTCATGACAGAAAGCGTGGGAGC 748  
609 CAACGGGAAAGCACTGTGCTGGGCGGACACTGACACTGAGAGCGAAAGCTAGGGAGC 550  
749 AAACAGATTAGATACCTGTGTAGTCCAGCCCTTAAACATGTCTCACTAGTTGGGCC 808  
549 GAATGGATTAGATACCCCACTAGTCTTACCCGTAACGATGATCTAGGCGCTGTGCG 490  
809 TTTATTGGCTTG--GTAAAGAGCTTACGCGTGAAGTTGACCGCTGGGAGTACGGTCG 866  
489 TATGACCCCGTGAAGTCTGTAGCTTACGCGTTTATGATATCCGCTGGGAGTACGTTGC 430  
867 CAAGATTAAACTCAAGAAATTGACCGGGAACCGGCAAGAGCGGTGATTATGTGATTA 926  
429 CAGAATGAATCAATCAAGAAATTGACCGGGAACCGGCAAGAGCGGTGATTATGTGATTA 370  
927 ATTGATGCAACCGGAAATCTTACTTACCTTGAATGTAGGAAATTTTCTTAGATATA 986  
369 ATTGATGCAAGAGGAACTTATCAAGGCTTGAATGCGGGAATCTCTTGAAGA 310  
987 GATTATGCG-CTTGGGGAAGCTTACAGAGTGTGAGGCTGTGCTGCTGCTGCTGCTGCT 1045  
309 GAGGGGTCCTTGGGGAAGCGGAGCAAGGTGTGATGCTGTGCTGCTGCTGCTGCTGCT 250  
1046 GAGATGTTGGGTTAAGTCCGCAAGAGCGCAACCTTGTCAATTAATTTGCGCATTTGG 1105

249 AAGGTTGGGTTAAGTCCGCAAGAGCGCAACCGCTGTGTATTAGTCCACCGTTGAG 190  
1106 TTGGGCACTTATAGACTCTCCGCTGACAAACCGGAGGAGTGGGATATACCTCACT 1165  
189 TTTGGAAACCTTGAACAGACTCCGGTGTATTAAGCCGAGGAAGGTGATGATGACCTCAAGT 130  
1166 CCTATGGCCCTTTATGGGCTTACAGGCTTACATATGAGCGCTTACAGAGGTTGC 1225  
129 CATCATGCCCCCTTATAGCCCTTGGGCGACACAGCTGTTCATATGCGCGGACAAAGGCTGC 70  
1226 CAACCCGAGAGGGAGCTAATCTCAGAAAGCGCTGTAGTCCGATCGAGTCTGCA 1285  
69 GATCCCGAGAGGGTGTAGCTTATCTCAAAAACCCGCTCTCAGTTGCGATTGACGCTGCA 10  
1286 CTCGACTCC 1294  
9 CTCGCTGC 1

RESULT 52  
BZ501502 831 bp DNA linear GSS 16-DEC-2002  
LOCUS  
DEFINITION  
BONDE57TR B0.1.6.2 KB tot Brassica oleracea genomic clone BONDE57,  
genomic survey sequence.  
ACCESSION  
BZ501502 GI:27018323  
VERSION  
BZ501502.1 GI:27018323  
KEYWORDS  
GSS.  
SOURCE  
Brassica oleracea  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica  
1 (bases 1 to 831)  
Ayele, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Uterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of *Brassica oleracea* and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490  
PUBMED  
Other\_GSSs: BONDE57TR  
Contact: Chris Town  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

## FEATURES

Location/Qualifiers  
1..831  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO100DH3"  
/db\_xref="taxon:3712"  
/clone="BONDE57"  
/clone\_1ib="BO.1.6.2 KB tot"  
/note="Vector: PHOS1; Site\_1: BstXI; 1.6-2 kb sheared  
total DNA inserted into PHOS1 using BstXI linkers"

## ORIGIN

Query Match 35.4%; Score 516.2; DB 9; Length 831;  
Best Local Similarity 78.1%; Pred. No. 7.7e-148;  
Matches 646; Conservative 0; Mismatches 178; Indels 3; Gaps 2;  
463 GAAGAGCAACCGGCTTACTACGTCGACAGACCGCGGTAATCGTAGGGTGAAGGCTTA 522  
5 GATATGACATCGGCTTACTCTGTGCGACAGCGCGGTATATACAGAGATGCAAGCGTTA 64  
523 ATCGGAATTACTGAGGCTAAAGGGTGCAGAGCGGCTTTGTAAGTGAATGTAATCC 582  
65 TCCGAAATGTTGGGCTAAAGCGTCTGTAGTGTGCTTTTAAAGTCCGCGCTCAAAATCC 124

QY 583 CGGGCTTAACCTGGGAATTGCGTTGAACTACAGGCTAGATGTCGAGAGGAGTGC 642  
DB 125 AGGGCTCAACCTGGACAGCGGTGAACTACCAAGCTTGAGTACGGTAGGGGAGAGG 184  
QY 643 GAATTCCTAGTGTAGCAGTGAATATGCGTGAATATGAAAGCAATCGATGCGAAGCA 702  
DB 185 GAATTTCCGGTGGACCGGTGAATGCTGAGATCGAAGAAACACCAACCGCGAAGAGA 244  
QY 703 GCTTCCTGGGTTAACATGACGCTCATGACAGAAAGCGTGGAGCAACACAGATTAGT 762  
DB 245 CTCTGCTGGGCGGACACTGACACTGAGAGACAAAGCTAGGGAGCGAATGGATTAGT 304  
QY 763 ACCCTGTTAGTCCAGGCTTAACGATGCACTAGTTTGGGCTTTTAGGCTTG-- 820  
DB 305 ACCCGAGTACTCTAGCCGTAAACGATGATCTAGGCCCTGTGCGTATCGACCCGTGA 364  
QY 821 GTACGAAAGCTAACCGGTGATGACGCTGCGGAGTACGCGTGCACAGATTAAACTC 880  
DB 365 GTGCTGTAGCTAACCGCTTAAGTATCCGCTGGGAGTACGTTGCGAAGATGAAACTC 424  
QY 881 AAAGAAATTGACGGGGACCGGCAACAGCGGTGATATGATGATTAACTGCAACGC 940  
DB 425 AAAGAAATTGACGGGGACCGGCAACAGCGGTGATATGATGATTAACTGCAACGC 484  
QY 941 GAAAAACCTTACTTACCTTGAATGATGACGAAATTTTCTAGATGATTAGT--CTTCG 999  
DB 485 GAAAGAACTTACCAAGGCTTGAATGACGCGGAATCTCTTGAAGAGAGGGGTGCTTCG 544  
QY 1000 GGAACGCTAACACAGGTGCTGATGAGCTGCTGACAGCTGCTGATGATGTTGGTTA 1059  
DB 545 GGAACGCGGACACAGGTGCTGATGAGCTGCTGACAGCTGCTGATGATGTTGGTTA 604  
QY 1060 AGTCCCGCAACGAGCGCAACCTTGTGATTAATGTCATGTTGGTGGCACTTTAT 1119  
DB 605 AGTCCCGCAACGAGCGCAACCTTGTGATTAATGTCATGTTGGTGGCACTTTAT 664  
QY 1120 GAGACTGCGCGGTGACAAACCGGAGAGAGTGGGATGATGCAAGCTTCATAGGCCCTTA 1179  
DB 665 CAGACTGCGCGGTGACAAACCGGAGAGAGTGGGATGATGCAAGCTTCATAGGCCCTTA 724  
QY 1180 TGGGTAGGGCTTACACGCTAATACATGAGCGGTACAGAGGGTGCACACCGCGAGGG 1239  
DB 725 TGCCCTGGGCGACACAGCTGCTACATGAGCGCGGACAAAGGGTCCGCGAGGGT 784  
QY 1240 GAGCTAATCTCAGAAAGCGGTGATGTCGAGTCCGATGCGAGTCTGCAAC 1286  
DB 785 GAGCTAATCTCAGAAAGCGGTGATGTCGAGTCCGATGCGAGTCTGCAAC 831

RESULT 53  
BH552149 835 bp DNA linear GSS 14-DEC-2001  
LOCUS BOHFF72TR BOHF Brassica oleracea genomic clone BOHFF72, genomic  
DEFINITION survey sequence.  
ACCESSION BH552149.1 GI:17803929  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Brassica oleracea  
Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta, eudicotyledons, core eudicotyledons;  
rosids; eustroids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE  
AUTHORS 1 (bases 1 to 835)  
Ayele, M., Haas, B. J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Utecher, T. R., Mortman, J. R., White, O. R. and Town, C. D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
JOURNAL PUBMED  
PUBMED 15805490  
COMMENT Other GSSs: BOHFF72TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.

FEATURES  
source  
location/Qualifiers  
1..835  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="T01000DH3"  
/db\_xref="taxon:3712"  
/clone="BOHFF72"  
/clone\_lib="BOHF"  
/note="Vector: pHOST1, Site 1: BexXI, 2-3 kb sheared  
genomic DNA inserted into pHOST using BexXI linkers"

ORIGIN  
Query Match 35.4%; Score 516.2; DB 9; Length 835;  
Best Local Similarity 77.7%; Pred. No. 7.7e-148;  
Matches 649; Conservative 0; Mismatches 183; Indels 3; Gaps 2;  
QY 479 ACTACGTCCAGCAGCCCGGTAAATACCTAGGTCAGACGCTTAATCGAAATTACTGCGC 538  
DB 1 ACTGTGTCCAGCAGCCCGGTAAATACAGAGATGCAAGCTTAATCGGAATGATTGGC 60  
QY 539 GTAAAGGTGGCAGGCGGCTTTGTAATGATGATGTAATATCCCGGCTTAACCTGGA 598  
DB 61 GTAAAGGTGGCAGGCGGCTTTGTAATGATGATGTAATATCCCGGCTTAACCTGGA 120  
QY 599 ATTGCGTTTAACTACAGGCTAGAGTGTGGCAGAGGAGTGAATTCATGCTGATGAC 658  
DB 121 CAGCGGTGAAATCTACCAAGCTTGAATGATGATGATGATGATGATGATGATGATGAT 180  
QY 659 AGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
DB 181 GTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
QY 719 CTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778  
DB 241 CTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
QY 779 CCTTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 836  
DB 301 CCTTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
QY 837 GTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 896  
DB 361 GTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 897 ACCCGCACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 956  
DB 421 GCCCGCACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 957 CCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1015  
DB 481 CCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
QY 1016 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1075  
DB 541 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
QY 1076 CAACCTTGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135  
DB 601 CAACCTTGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
QY 1136 AACCGGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1195  
DB 661 AACCGGAGAGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
QY 1196 CGTAAATCAATGCGCGGTACAGAGGTTGCCAACCGCGGAGGAGAGTAACTTCAGAA 1255  
DB 721 CGTAAATCAATGCGCGGTACAGAGGTTGCCAACCGCGGAGGAGAGTAACTTCAGAA 780

QY 1256 GCGCGTCGTAGTCCGAGTCGAGCTGCAACTGCTCCGTGAAGTGGAAATCC 1310  
DB 781 CCGCTCTCAGATTCCGATTGCGAGGCTGCAACTGCTCGATGAAAGCCGAATCGC 835

RESULT 54  
CM001852 1050 bp DNA linear GSS 23-SEP-2004  
LOCUS ZMBHhg0002h20.f ZMBHhg Zea mays genomic clone ZMBHhg0002h20 5',  
DEFINITION genomic survey sequence.  
ACCESSION CM001852  
VERSION CM001852.1 GI:52576905  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 1050)  
Ma,J., Sami,guel,P., Liu,R., Haller,K., Soderlund,C. and  
Bennetzen,J.  
REFERENCE ZMBH sequences  
AUTHORS Unpublished (2004)  
JOURNAL Contact: Jeff Bennetzen  
COMMENT Bennetzen Lab  
The University of Georgia  
Department of Genetics, C426a Life Sciences Building, Athens, GA  
30602, USA  
Tel: 706-542-3698  
Fax: 706-583-0972  
Email: maize@uga.edu  
Plate: 0002 row: h column: 20  
Class: BAC ends.

FEATURES  
source  
Location/Qualifiers  
1..1050  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/culivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMBHhg0002h20"  
/tissue\_type="immature ear"  
/dev\_stage="6-8 weeks"  
/lab\_host="DH10B"  
/clone\_1lb="ZMBHhg"  
/note="Vector: TOPopcr4; Site\_1: EcoRI; Site\_2: EcoRI"

ORIGIN  
Query Match 35.4%; Score 515.6; DB 10; Length 1050;  
Best Local Similarity 75.4%; Pred. No. 1.3e-147;  
Matches 706; Conservative 0; Mismatches 224; Indels 6; Gaps 5;

QY 527 GAATTACTGCGCGTAAGGTCGCGAGCGCGCTTTGTAAGTCAGATGTAATCCCGCG 586  
DB 2 GGAATGATTGGCGGTAAAGCGTCTGTAGTGGCTTTTCAAGTCCCGCTCAATCCGAGG 61

QY 587 CTTAACTGGGAATTGGCTTTGAACTCAAGCTAGAGTGGCAGAGGAGGTGAAT 646  
DB 62 CTCAACCTCGACAGCGGTGGAACCTCAAGCTGAGTACGTTAGGGGAGGGAAT 121

QY 647 TCCATGTGTAGCAGTGAATGCGTAGAGATGTGAAGAACATCGATGGCGAGCACCT 706  
DB 122 TTCGGGTGAGCGGTGAAATGCACTTGAATCGAAAGAACCAACGCGGAGAACACTCT 181

QY 707 CTTGGGTTAACTGACGCTCTGACGAGAAAGCGTGGGAGCAAAAGAGATTAGATCC 766  
DB 182 GCTGGGCGGACACTGACACTGAGAGAGCAAGAGCTAGGGGAGCAATATGGATTAGAGCC 241

QY 767 TGTAGTCCAGCGCCTTAAGAGATGCACTAGTGTGGGCTTATTAGGCTTG--GTAA 824  
DB 242 CAGTAGTCTTAGCCGTAAAGCATGATGATCTAGGTGCTGTGCACTCGACCTGTCAATGC 301

QY 825 CGAAGCTAACGCGTGAAGTTGACCGCTGGGAGATACGCTGCAAGATTAAACTCAAG 884

DB 302 TGTACTTAACCGCTTAAGTATCCCGCTGGGAGTACGTTGCAAGTAAGAACTCAAG 361  
QY 885 GAATTGACGGGAGCCCGACAAAGCGGTGATTTATGTGATTAATTCAGTCAACCGGAA 944  
DB 362 GAATTGACGGGAGCCCGACAAAGCGGTGATTTATGTGATTAATTCAGTCAACCGGAA 421

QY 945 AACCTTACCTTACCTTGAACATGCTAGCGAATTTTCAGAGATGATTAGTG--CTTGGGGA 1003  
DB 422 AACCTTACCGAGGCTTGAACATGCTAGCGAATTTTCAGAGATGATTAGTG--CTTGGGGA 481

QY 1004 CGCTAACACAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063  
DB 482 CGCGACACAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541

QY 1064 CCGCAACGAGCGCAACCTTGTCTATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123  
DB 542 TCGCAACGAGCGCAACCTTGTCTATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 1124 TCGCGGTGACAAACCGGAGAAAGTGGGATGACGTGAAGTCTCATGCGCCCTATAGGG 1183  
DB 601 CCGCGGTGTTAAGCCGAGAGAAAGAGATGAGGCCAAATCATCATGCCCTTATAGCC 660

QY 1184 TAGGGCTTCAACGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1243  
DB 661 CTGGCGACACACGTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720

QY 1244 TAATCTCAGAAAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1303  
DB 721 TAATCTCAGAAAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY 1304 GAATGCTGATTAATCGCGGATCAG--CATGCGCGGTGAATTAATTAATTAATTAATTAAT 1362  
DB 781 GAATGCTGATTAATCGCGGATCAG--CATGCGCGGTGAATTAATTAATTAATTAATTAAT 840

QY 1363 CACCGCGGTGACACCATGAGGAGTGGGTTTCAACAGAACAGTAGTC--TAACTGTAAG 1421  
DB 841 CACCGCGGTGACACCATGAGGAGTGGGTTTCAACAGAACAGTAGTC--TAACTGTAAG 900

QY 1422 AGGCGCTTGCACAGTGTGATTCATGATGAGGAGT 1457  
DB 901 AGGCGGATGCTTAAGGCTAGGCTTGGCATGAGAGTG 936

RESULT 55  
BH578853 857 bp DNA linear GSS 14-DEC-2001  
LOCUS BOHQE24TR BOHQ Brassica oleracea genomic clone BOHQE24, genomic  
DEFINITION survey sequence.  
ACCESSION BH578853  
VERSION BH578853.1 GI:17831134  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 857)  
Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490  
Other\_GSSs: BOHQE24TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR

Class: sheared ends.

FEATURES	Location/Qualifiers
source	1. .857

```

organism="Brassicica oleracea"
mol_type="genomic DNA"
strain="Tol000DH3"
db_xref="taxon:3712"
clone="BOHQE24"
clone_1b="BOHQ"
name="vector: pHS01, site_1: BstXI, 2-3 kb sheared
genomic DNA inserted into pHS01 using BstXI linkers

```

**ORIGIN**

Query Match	35.3%	Score 515	DB 9	Length 857
-------------	-------	-----------	------	------------

Best Local Similarity 77.8%; Pred. No. 1.8e-147;  
Matches 670; Conservative 0; Mismatches 165; Indels 26; Gaps 3

Matches 670; Conservative 0; Mismatches 165; Indels 26; Gaps 3

QY	228	GATAAAGCCCTAACCAAGGCGAGCATCAGTAGTTGGCTCTAGAGAGACGACCACTCACTG	287
Db	857	GGCAATAGCTTACCAAGGCGATGATCAGTAGCTGTTCCGAGAGGATGATCAGCACTCACTG	798
QY	288	GGACTGAGACACGCGCCAGACTCTCTACGGGAGGACAGCACTGGGGAAATTTTGGACATGGG	347
Db	797	GGACTGAGACACGCGCCCAACTCTCTACGGGAGGACAGCACTGGGGAAATTTTCCGCAATGGG	738
QY	348	CGCAAGCCCTGATCCAGCAATGCGCGGTGAGTGAAGAAAGGCTTCGGGTGTGTAAGCTCTT	407
Db	737	CGAAAGCTTGACGAGACATATGCCGCTGAGAGGTAGAAAGGCTCTACGGGTCTTGAACTTCTT	678
QY	408	TCAGTCGAGAGAGAAAAGTTACGGTAAATATCTGATCATGACGGTATCGACAGAGAA	467
Db	677	TTCCCAAGAGAA-----AGCAATGACGGTATCTGGGAAATA	641
QY	488	AGCACCGGCTTACTATGTCGACAGACCGCCGGTAAATACGTAGGGTCCAGCGCTTAATCGG	527
Db	640	AGCATCCGGTAACTCTGTGCCAGACGCGCGGTAATACAGAGATGCAAGCGTTATCCGG	581
QY	528	AATTACTGGCGCTTAAAGGTGCGACGCGGCTTGTGTAATCAATGATGAAATCCCGGCGC	587
Db	580	AATGATTTGGCGCTTAAAGCGTCTGTAGTGGCTTTTAACTCCGCGCTCAAAATCCAGGGC	521
QY	588	TTAACTTGGGAATTGCGTTTGAATCTCAAGGCTTAGAGTGTGGCAGAGGAGGTGGAAAT	647
Db	520	TCAACCTCGGACAGGGGGGTGGAATCTACCAAGCTTGAGTACGGTAGGGCAGAGGAAAT	461
QY	648	CCATGTGTAGCAGTGAAATGCGTAGAGATATGGAGAAACATGATGATGGCCAGACGACCTTC	707
Db	460	TCCGGGTGGACGCGTAAATGCGTAGAGATCGAAGAAACCAACGCGCCAAAGCACTCG	401
QY	708	CTGGGTTAACTATGACGCTCATGACAGAAAGGTGGGGAGCAAAACAGATTAGATACCT	767
Db	400	CTGGGCGGACACTGACACTGAGAGACGAAAGCTAGGGGAGCGAATGGGATTAGATACCCC	341
QY	768	GGTAGTCCACGCGCCTTAAACGATGTCAACTAGTTGTGGGCTTATTAAGCTTG--GTAAAC	825
Db	340	AGTAGTCTCTAGCGCGTAAACGATGATATACAGGCGCTGTGCTATCGAACCCGTCAGTGCT	281
QY	826	GAAAGCTAACGCGTGAAGTTGACACCGCTCGGGGAGTACGCTCGCAAGATTAAATCTCAAG	885
Db	280	GTAGTAAACGCGTTTAAAGTATCCCGCTGGGAGGTAGTTTCGCAAGATATAAACTCAAGG	221
QY	886	AATTACGGGGGACCGGCAACAGCGGTGATTTATGTGATTTAATTCGATCCAAACGCAAAA	945
Db	220	AATTACCGGGGCGCGGACAAACGCGGTGAGCATGTGGTTTAATTCGATCCAAACGCAAAA	161
QY	946	ACCTTACCTAACCTTTGACATGTAGCGAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAC	1000
Db	160	ACCTTACCAAGGCGCTTGACATGCCGCAATCTCTTAAAGAAGGGGCTCCTTCGGGAAC	101
QY	1005	GCTAACACAGGTGCTGCAATGCTGTCTCAAGCTCGTGTGTGAGATGTTGGGTAAAGTCC	106
Db	100	GGGACACAGGTGCTGCAATGCTGTCTCAAGCTCGTGTGTGAGATGTTGGGTAAAGTCC	41

```
QY      1065 CGCAACGAGCGCAACCTTGT   1085
          |||||
Db      40  CGCAACGAGCGCAACCTTGT   20
```

RESULT	56				
LOCUS	BH704946/c				
DEFINITION	BH704946	835 bp	DNA	linear	GSS 20-FEB-2002
ACCESSION	BH704946				
VERSION	BH704946				
KEYWORDS	BH704946.1	GI:18787033			
SOURCE	GSS				
ORGANISM	Brassica oleracea				
	Brassica oleracea				
	Brassica oleracea				

FEATURES	Location/Qualifiers
source	1. .835

```
'organism="Brassicica Oleracea"
mol_type="genomic DNA"
strain="T01000DH3"
db_xref="taxon:3712"
clone="BOM1Q71"
clone_1b="BO_2_3_KX"
note="Vector: pGHS1, Site_1: BstXI, 2-3 kb sheared
genomic DNA inserted into pGHS1 using BstXI linkers
```

## ORIGIN

Query Match	Score	DB	Length
35.3%	514.8	9	835

Best Local Similarity 78.2%; Pred. No. 2.1e-147;  
Matches 644; Conservative 0; Mismatches 177; Indels 3; Gaps 2;

Matches 644; Conservative 0; Mismatches 177; Indels 3; Gaps 2;

Oy	AGGGTGCACAGGCTTAATTCGAAATTACTGGGCGTAAAGGATCCGACGCGCTTTGTAGT	508
Db	AGGATGCAGCGTTATCCGGATATTTGGGCGTAAAGCGTCTGTAGTGGCTTTTAAAGT	833
Oy	CAGATGTGAATTCGCCGGGCTTAACTGGGAATTGCGTTTGAACCTAACAGGCTAGAGTG	568
Db	CCGCGCTGAATCCGAGGCGTCAACCGCTGACAGGCGGCTGGAACTACAGAGCTTGAGTA	773
Oy	TGGCAGAGGGAGGTGGAAATTCATGTGTGACAGTGAATAATGCTGTAGATATGGAAGAAC	628
Db	CGGTAGGGGGCAGAGGGAATTTCCGCTGGAGGGGTGAATGTGTAGAGTTCGGAAGAACA	713
Oy	TGGATGGGGAAGGAGCGCTCCTGGGTTAACTACAGCTCATGACAGGAACGTGGGGAG	688
Db	CCAAACGGGGAAGACCTTGCTGGGCGGCACACTACCTGAGAGACGAAGACTAGGGGAG	653
Oy	CAAAACAGATTAGATACCCTGTGTAGTCCACGCGCTTAACAGATGTCAACTAGTTTGGGC	748
Db	CGAATGGGATTAGATACCCACGATAGTCTTACGCCGTAAACGATGGATCTACGCGCTGTGC	593
Oy	CTTATTAGGCTTG--GTACGAGACTAACGCGTGAAGTTGACCGGCTGGGGAGTACGGTC	808

Db 533 GATATGACCCGTCGACGTGCTGTAAGCTTAACCGCTTAAGTATCCCGCTGGGAGTACGTTTC 474

QY 866 GCAAGATTAAAACCTCAAGAAATTGACGGGGAACCCGACAAAGCGGTGATTAATGATTT 925

Db 473 GCAAGATTAAAACCTCAAGAAATTGACGGGGAACCCGACAAAGCGGTGATTAATGATTT 414

QY 926 AATTGATGCAACGGCAAAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTA 985

Db 413 AATTGATGCAACGGCAAAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTA 354

QY 986 AATATTAGTTCCTTGGGGAACGGTCAACGGTCTGATGCGTGTGATGCGTGTGATGCGTGTG 1044

Db 353 AAGAGGGGTCTCTTGGGGAACGGTCAACGGTCTGATGCGTGTGATGCGTGTGATGCGTGTG 294

QY 1045 TGAGATGTTGGGTTAAGTCCCGCAACGACGCAACCCCTTGTCAATTAATTTGCAATTTG 1104

Db 293 TAAAGGTGTTGGGTTAAGTCCCGCAACGACGCAACCCCTGTTTGTGTTGCGACCGTTGA 234

QY 1105 GTTGGGACCTTAATGACATGCGGTGACAAACCGAGAGAGTGGGATGACGTCAAG 1164

Db 233 GTTGGGACCTTAATGACATGCGGTGACAAACCGAGAGAGTGGGATGACGTCAAG 174

QY 1165 TCTCATGCGCCCTTAATGAGGCTTTCACAGTAAATGACATGCGGTGACAGAGGTTG 1224

Db 173 TCATCATGCGCCCTTAATGAGGCTTTCACAGTAAATGACATGCGGTGACAGAGGTTG 114

QY 1225 CCAACCCGAGGAGGAGTATCTCAAGAAAGCGCGTGTGATGCGGTGATGCGGTGATGCG 1284

Db 113 CGATCCCGAGGAGTATCTCAAGAAAGCGCGTGTGATGCGGTGATGCGGTGATGCG 54

QY 1285 ACTCGACTCGGTAAGTCGGAATCGTAGTAATCGCGGATGACG 1328

Db 53 ACTCGCTGATGAGCGCGAATCGCTAATGATCGCGGTGACG 10

RESULT 57  
BH473870/c 845 bp DNA linear GSS 13-DEC-2001  
LOCUS BOHGQ277R BOHG Brassica oleracea genomic clone BOHGQ27, genomic  
DEFINITION survey sequence.

ACCESSION BH473870  
VERSION BH473870.1 GI:17681981  
KEYWORDS GSS.

SOURCE  
ORGANISM Brassica oleracea

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

TITLE 1 (bases 1 to 845)  
JOURNAL Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
PUBMED Uteback, T.R., Mortman, J.R., White, O.R. and Town, C.D.  
Genome Res. 15 (4), 487-495 (2005)  
COMMENT Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis

Other GSSs: BOHGQ27TF  
Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

FEATURES  
Location/Qualifiers

1..845  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOHGQ27"  
/clone\_id="BOHG"

ORIGIN  
Query Match 35.3%; Score 514.8; DB 9; Length 845;  
Best Local Similarity 78.0%; Pred. No. 2.1e-147;  
Matches 658; Conservative 0; Mismatches 182; Indels 4; Gaps 3;

QY 588 TTAACCTGGGAATTCCTTTGAAACTACAGGCTAGAGTGTGACAGAGGAGTGAATT 647  
Db 845 TCAACCTGGGACAGCGGTGGAACCTACCACTGAAGTACCTAGAGGAGGAGGATTT 786

QY 648 CCATGTGTAGCAGTGAATGCGTAGAGATATGAAAGAACATGATGCGAGCAGCCTTC 707  
Db 785 CCGGTTGAGCGGTGAATGCTTAAGATCGAAAGAAACCAACGCGGAAAGCACTCTG 726

QY 708 CTGGGTTAACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTGATTCCT 767  
Db 725 CTGGGCGCACACTGACACTGAGAGACGAAAGCTAGGGAGCGAATGGGATTAATGATCC 666

QY 768 GGTATGCAAGCCCTTAACGATGTCACTAGTTGTTGGCCCTTATTAGGCTTG -GTAC 825  
Db 665 AGTAGTCTTAGCCGTAAAGATGATTAAGGCGCTGTGCTGATGCAACCGTGCAGTGT 606

QY 826 GAAAGCTAAGCGGTAAGTGAACCGCTGGGAGTACGCTGCGAAGTAAACTCAAG 885  
Db 605 GTAGCTTAAGCGCTTAAGTATCCCGCTGGGAGTACGTTTCGAAAGATGAACCTCAAG 546

QY 886 AATTGACGGGAGACCCGACAAAGCGGTGATTAATGATTAATTCGATGCAACGCAAAA 945  
Db 545 AATTGACGGGAGCCGACAAAGCGGTGATTAATGATTAATTCGATGCAACGCAAGA 486

QY 946 ACCTTAACCTTACCTTACATGTAAGCAATTTTCTAAGATTAATTTGTC -CTTGGGAAC 1004  
Db 485 ACCTTAACCAAGGCTTGAATGCGCGCAATCTCTTGAAGAGAGGAGTGTGCGGAAC 426

QY 1005 GCTAACACAGGTGCTGATGCGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1064  
Db 425 GCGAACAAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366

QY 1065 CGCAACGAGCGCAACCTTGTCAATTAATTCGCAATTTGTTGGSCACTTAAATGAG 1124  
Db 365 CGCAACGAGCGCAACCTTGTCAATTAATTCGCAATTTGTTGGSCACTTAAATGAG 306

QY 1125 TCCCGGTGCAAAACCGAGAGAGTGGGATGACGTCAGTCTCTATGCGCTTATGGGT 1184  
Db 305 TCCCGGTGATTAAGCCGAGAGAGTGAAGTGAATGATGCTCAATGCTCTTATGCCC 246

QY 1185 AGGCTTTCACAGTAAATGACATGCGCGCTACAGAGGAGTGGCCAAACCGGAGGAGGT 1244  
Db 245 TGGGCGACACAGTGTACATGCGCGGAGCAAAAGGAGTGGCATCCCGGAGGAGTGA 186

QY 1245 AATCTGAAAGAGCGGTGATGCGGATGCGAGTCTGCAACTGCACTCGGTGAAGTCGG 1304  
Db 185 AATCTGAAAGAGCGGTGATGCGGATGCGAGTCTGCAACTGCACTCGGTGAAGTCGG 126

QY 1305 AATCGTAGTAATGCGGAGTCAAG -CATGTGCGGTGAATGCTTCCGGGTCTTTGTAAC 1363  
Db 125 AATCGTAGTAATGCGGAGTCAAG -CATGTGCGGTGAATGCTTCCGGGTCTTTGTAAC 66

QY 1364 ACCGCGGTGACATGAGAGTGGGTTTCAACGAAGAGTGAAGTGAACCGTGAAGAG 1423  
Db 65 ACCGCGGTGACATGAGAGTGGGTTTCAACGAAGAGTGAAGTGAACCGTGAAGAG 6

QY 1424 GGCG 1427  
Db 5 GGCG 2

RESULT 58  
DR027023 895 bp mRNA linear EST 26-MAY-2005  
LOCUS Osmo01058 F. cylindrus osmotic stress library Frag11ar10p51  
DEFINITION



## ORIGIN

/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONR251"  
/note="lib="BO.1.6.2 KB tot"  
/note="Vector: pHO51, Site 1: BstXI, 1.6-2 kb sheared  
total DNA inserted into pHO51 using BstXI linkers"

Query Match 35.3%; Score 513.6; DB 9; Length 862;  
Best Local Similarity 78.0%; Pred. No. 5e-147;  
Matches 644; Conservative 0; Mismatches 179; Indels 3; Gaps 2;

448 ATGACGGTATCTGGGAATTAAGCATCGGCTAACTCTGTGCGACAGCCGGTAAATACAG 767  
826 ATGACGGTATCTGGGAATTAAGCATCGGCTAACTCTGTGCGACAGCCGGTAAATACAG 767  
508 AGGCTGCAAGCCCTTAATCGGAATTAATCGGCGGCTAAAGGGGCGGCGGCTTGTAACT 567  
766 AGGATCAAGCCCTTAATCGGAATTAATCGGCGGCTAAAGGGGCGGCTTGTAACT 707  
568 CAGATGGAATCCCGGCTTAACCTGGGAATTTGGTTGAATCAAGGCTAGAGTG 627  
706 CGCCCTCAATCCAGGCTCAACCTGACAGGGGCTGGAATCAAGCTTAGTA 647  
628 TGGCAGAGGAGGTGAATTCATGTGTAGCAAGTGAATTCGTAGAGATTTGGAAGACA 687  
646 CGGTAGGGGAGAGGGAATTTCCGGTGGAGCGGTGAATTCGTAGAGATTCGGAAGACA 587  
688 TCGATGCGCAAGGACGCTCTCTGGTTAACTAGACGCTCATGACGAAAGCGTGGGAG 747  
586 CCAACGCGCAAGCACTCTGCTGCTGCGGACAGTGAACCTAGAGACCAAGGCTAGGGAG 527  
748 CAAACGAGATTAGATACCTGTAGTCCAGCCCTTAACGATGTCACTAGTTGTTGGC 807  
526 CGAATGGATTAATGATACCCAGATAGTCTTAGCCGTAAACATGATATCTAGGGCGTGTGC 467  
808 CTATTAGGCTTG--GTACGAAGCTTAACGCTGGAATTGACCGCTGGGAGTACGCTC 865  
466 GTATCGACCCGTCAGTGTAGCTTAACGCTGAATATCCGCTGGGAGTACGCTTC 407  
866 GCAAGATTAATCAAGGAATTAAGCGGAGACCCGACAAAGGGGTGATTAATGATG 925  
406 GCAAGATTAATCAAGGAATTAAGCGGAGACCCGACAAAGGGGTGATTAATGATG 347  
926 AATTGATGCAAGCGGAAACCTTACCTACCTTGCATGATGAGCAATTTTCTAGAGAT 985  
346 AATTGATGCAAGCGGAAACCTTACCTACCTTGCATGATGAGCAATTTTCTAGAGAT 287  
986 AGATTAGTG-CTTCGGGAACGCTTAACACAGGTGCTGATGCTGTCAAGCTGTGTCG 1044  
286 AGAGGGGTGCTTCGGGAACGCGGACACAGGTGCTGTGCTGTCAAGCTGTGTCG 227  
1045 TGAATGTTGGGTTAATCCCGCAAGAGGCAACCTTGTCTAATTAATTCGATCATTTG 1104  
226 TAAGTGTGTGGTTAATTCCTCGCAACAGAGGCAACCTTGTGTATGTGCAACCGTTGA 167  
1105 GTTGGGCACTTTAATGAGACTGCGCGTGAACAAACCGGAGGAAGGTGGGATGAGCTCAAG 1164  
166 GTTTGGAACCTTGAACAGACTGCGCGTGAATAGCGGAGGAAGGTGAGATGAGCTCAAG 107  
1165 TCCTCATGAGCCCTTAATGGGTTAGGGCTTCAACGTAATTAATGAGCGGTGACAGGGTTG 1224  
106 TCATCATGAGCCCTTAATGGGTTAGGGCTTCAACGTAATTAATGAGCGGTGACAGGGTTG 47  
1225 CCAACCGGAGAGGAGGTGAATTCAGAAAGCGCGTGTAGTACG 1270  
46 CGATCCGCGAGGGTGAATTAATCAAAAACCGTCTCAAGTTG 1

RESULT 60  
CV291378/c CV291378 837 bp mRNA linear EST 08-JUN-2005  
LOCUS aof01-1ms3-g08 aof01 Asparagus officinalis cDNA clone

accession aof01-1ms3-g08 5', mRNA sequence.  
CV291378  
VERSION CV291378.1 GI:52577614  
KEYWORDS EST.  
SOURCE Asparagus officinalis (garden asparagus)  
ORGANISM Asparagus officinalis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;  
Asparagaceae; Asparagus.

REFERENCE 1 (bases 1 to 837)  
dePamphilis,C., Carlson,J., Ma,H., Solits,D., Solits,P.,  
Oppenheimer,D., Fronlich,M., Doyle,J., Tankley,S., Webb,M.,  
Lebens-Mack,J., Landherr,L., Iluv,D. and Wall,K.  
Generation of ESTs from early male inflorescences of Asparagus  
officinalis  
Unpublished (2004)

TITLE JOURNAL  
COMMENT Contact: Claude dePamphilis or James Lebens-Mack  
Muehler Laboratory  
Penn State University  
208 Muehler Laboratory, Department of Biology, ATTN Rm212, Penn  
State University, University Park, PA 16802, USA  
Tel: 814 863 6413  
Fax: 814 865 9131  
Email: cwd3@psu.edu or jhl10@psu.edu

FEATURES  
source The sequence provided is trimmed of vector and low quality regions.  
Full sequence and original trace file are available from the Plant  
Genome Network website (http://pgn.cornell.edu)  
Plate: aof01-1ms3 row: 9 column: 08  
Seq primer: M13F.  
Location/Qualifiers

1..837  
/organism="Asparagus officinalis"  
/mol\_type="mRNA"  
/db\_xref="PGN:aof01-1ms3-g08"  
/db\_xref="taxon:4686"  
/clone="aof01-1ms3-g08"  
/ribose\_type="male inflorescences"  
/lab\_host="SOLR"  
/note="Vector: pBluescript SK (+/-); Site 1: BstXI;  
/note="lib="aof01"  
/note="Vector: pBluescript SK (+/-); Site 1: BstXI;  
Site 2: XhoI; This is a directionally cloned,  
non-normalized library. This library has been generated by  
the Floral Genome Project (FGP). The Floral Genome Project  
is funded by NSF's Plant Genome Research Program  
(DBI-0115684). More information about the project can be  
obtained at http://fgp.bio.psu.edu"

## ORIGIN

Query Match 35.2%; Score 513; DB 7; Length 837;  
Best Local Similarity 77.8%; Pred. No. 7.6e-147;  
Matches 668; Conservative 0; Mismatches 165; Indels 26; Gaps 3;

221 TTGGTGGGTTAAAGGCTTAACAGCGGACGATGATGTTGTTCTGAGAGACGACCGAG 280  
837 TTGGTGGGTTAAAGGCTTAACAGCGGACGATGATGATGTTGTTCTGAGAGAGATGATCAGC 778  
777 CACACTGGGACTAGACACGGCCAGACTCTACCGGAGGACACAGTGGGAAATTTTCCG 718  
281 CACACTGGGACTAGACACGGCCAGACTCTACCGGAGGACACAGTGGGAAATTTTGA 340  
341 CAATGGGCGCAACCTGATCCAGCAATGCCGCTGAGTGAAGAGGCTTCGGGTTGTA 400  
717 CAATGGGCGCAACCTGATCCAGCAATGCCGCTGAGTGAAGAGGCTTCGGGTTGTA 658  
401 AGCTCTTCACTGAGAGAAAGTTACCGTAAATATGATGATCATGACGTTATGCA 460  
657 ACTTCTTTTCTCAGAGAGAG-----AGCAATACCGTATCTG 621  
461 CAGAGAGAGACCGGCTTAATCACTGTCACAGACCGCGGTAATACGTAAGGTGCAAGCT 520  
620 AGGAATTAAGCATCGGCTTAATCTGTGTCAGACACCGCGGTAACAGAGATGCAAGCGT 561  
521 TAATCGAATTACTGGCGTTAAAGGGTGGCGGCGCTTGTAAAGTCAATGTGAATATC 580



560 TATCCGGAATGATTTGGGGCTAAAGGCTGTGATGGCTTTTCAAGTCCGGCGTCAATC 501

Qy 581 CCCGGGCTTAACCTGGGAATTGCGTTTAAACTACAGGCTTGAATGTGGCAAGGAGG 640

Db 500 CCAGGGCTCAACCTTGACAGGCGGTGAAAACTACAGCTGGAGTACGGTAGGGGCGAGA 441

Qy 641 TGGAAATTCATGTGTGACAGTGAATATGGTAGATATGAAAGAACATCGATGGCGAAG 700

Db 440 GGGAAATTTCCGGTGGAGCGGTGAATATGGTAGAATTCGGGAAGAACACCAACGGGGAAG 381

Qy 701 CAGCCTCTGGGTTAACACTGACGCTACGCAGAAAGCTGGGGAGCAACAGGATTAG 760

Db 380 CACTCTGTGGGCGCAGCACTGACACTGAGAGCGAAACCTAGGGGAGCAAAATGGGATTAG 321

Qy 761 ATACCCTGTGTGTCCAGGCGCTAAACGATGTCATTAATTGTGGGCTTATTTAGGCTGT 820

Db 320 ATACCCAGTAGTCTTAAGCCGTTAAACGATGATTAATTAGGCGCTGTGCGTATCGAACCGGTG 261

Qy 821 --GTAAACGAAGCTAACGCGTGAAGTTGAACGCGCTTGGGGAGTAGCGGTGCAAGATTAAAC 878

Db 260 CAGTGTGTAGCTAACGCGTTAAGTATCCGCTTGGGAGTAGTCTTCCAGAAATGAAC 201

Qy 879 TCAAAGGAATTGAACGGGGACCCCGCACAGCGGTGATTAATGTGATTAATTCGATGCAC 938

Db 200 TCAAAGGAATTGAACGGGGACCCCGCACAGCGGTGAGCATGTGTTTAATTCGATGCACA 141

Qy 939 GCGAAAAAACCCTTACCTTGAACATGTAGAGCAATTTTCTAGAGATAGATTAGT-CTT 997

Db 140 GCGAAGAACCTTACCAAGGCTTGAACATCCCGGAATCCTCTTGAAGAAAGAGGGGTGCTT 81

Qy 998 CCGGAACGCTAACACAGGTGCTGCATGGCTGTGTCAGCTGTGTGCTGTAAGATTGGGT 1057

Db 80 CCGGAACCGGAGCACAGGTGTGTGATGCTGTGTCAGCTGTGTCGCTGTAAGATTGTGGGT 21

Qy 1058 TAAGTCCCGCAACGAGCGC 1076

Db 20 TAAGTCCCGCAACGAGCGC 2

RESULT 61  
B2446240/c  
LOCUS B2446240 822 bp DNA linear GSS 13-DEC-2002  
DEFINITION BOMJ104TR BO.1.6.2\_KB\_tot Brassica oleracea genomic clone BOMJ104,  
genomic survey sequence.  
ACCESSION B2446240  
VERSION B2446240.1 GI:26709961  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 822)  
AUTHORS Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
Ureback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
TITLE Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
JOURNAL Genome Res. 15 (4), 487-495 (2005)  
PUBMED 15805490  
COMMENT Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.  
location/Qualifiers  
    1..822  
    /organism="Brassica oleracea"  
    /mol\_type="genomic DNA"

```

/strain="TO1000DH3"
/db xref="taxon:3712"
/clone="BONJ104"
/clone_1lb="BO_1.6.2 KB tot"
/note="Vector: pHO1, Site 1: BstXI, 1.6-2 kb sheared
total DNA inserted into pHO1 using BstXI linkers"

```

Query Match	35.2%	Score 512.8	DB 9	Length 822
Best Local Similarity	78.1%	Pred. No. 8.7e-147		
Matches 642	Conservative 0	Mismatches 177	Indels 3	Gaps 2
QY	455	TATGACAGAAAGAACACCGGCTACTACTGTCACGACGACCGCGGTAACTACTGAGCTGC	514	
Db	822	TATCTGGGAAATAAAGCATCGGCTAACTCTGTGCCAGACGCGCGGTATATACAGAGATGC	763	
QY	515	AAGGTTAATCCGAATTACTGGGGGTAAAGGGTGGCGAGGGCGGCTTTGTAGTCAGATGT	574	
Db	762	AAGGTTATCCGGAGATATATGGGGTAAAGGCGTCTGTAGGTGGCTTTTAAAGTCCGCGT	703	
QY	575	GAATCCCCCGGGCTTAACTGGGAATTCGTTTGAACCTACAGGCTAAGATGTGGCAGA	634	
Db	702	CAATCCCAAGGCGTCAACCTCGGACAGGCGGTGAACCTACAGGCTTGAATACGATAG	643	
QY	635	GGGAGGTGAATTCATGTGTAGCATGTGAATATGCTGTAGATATGAAAGAACATCGATGG	694	
Db	642	GGCGAGAGGAATTTCCGCTGAGCGGTGAATATGCTAGATATCGAAAGAACACCAACGG	583	
QY	695	CGAAGGACGCTCCCTGGGTTAAACAATGACGCTACGACAGAAACGTTGGGAGCAACAG	754	
Db	582	CGAAGACACTGTGTTGGGCGGACACTACACTAGAGAACGAAACCTTAGGGAGCGAATGG	523	
QY	755	GATTAGATACCCCTGGTATGCCACGCGCTTAAACGATGTCAACTAAGTTGTGGGCTTATTA	814	
Db	522	GATTAGATACCCAGTATGCTCTAGCCGTAAACGATGATCTAGCGCGCTGTGCGATGCA	463	
QY	815	GCGCTTG--GTACGAACTAACCGGTGAAGTTGACCGCTGGGAGATACGCTCCCAAGT	872	
Db	462	CCCGTGAAGCTGTACTACTAACCGGTTAAGTATCCCGCTGGGAGATACGTTCCCAAGAA	403	
QY	873	TAAACCTCAAGGAATGACGGGGACCCGACAGACGCGTGGATTAATGATGATTAATTGCA	932	
Db	402	TGAACCTCAAGGAATGACGGGGACCCGACAGACGCGTGGAGCATGTGGTTAAATTGCA	343	
QY	933	TGCAACCGGAAAACTTACTCTACCTTGAATGATAGCGAATTTTCTAGAGATGATTAAG	992	
Db	342	TGCAACCGGAAAACTTACCAAGGCTTGAATGATGATGATGATGATGATGATGATGATG	283	
QY	993	TG--CTTGGGAGCGCTTAACACAGGTGCTGCATGCTGTCTGCTGCTGTCTGTGATGATG	1051	
Db	282	TGCTTCTGGGAGCGCGGACACAGGTGTGTGATGCTGTCTGTGCTGTGCTGTGATGATG	223	
QY	1052	TTGGGTTAAGTCCCGCAACGAGCGCAACCTGTGCATTAATTGGCATCATTTGGTTGGCG	1111	
Db	222	TTGGGTTAAGTCCCGCAACGAGCGCAACCTGTGTGTTAAGTGTGCAACGTTGAATTTGGA	163	
QY	1112	ACTTTAATGAGACTGCGGATGACAAACCGAGAGAGGTGGAGATGACGTCAAGTCTTCAT	1171	
Db	162	ACCTGAAACAGACTGCGGATGATTAAGCCGAGAGAGGTGAGATGAGTCAAGTCATCAT	103	
QY	1172	GGCGCTTATAGGATGGGCTTACACCGTAAATACATATGCGCGCTAACAGAGGTGTCACACC	1231	
Db	102	GGCGCTTATATCCCTGTGGGCGACACACGTTGATCAATATGACCGGAGCAAAAGGTCGCGATCC	43	
QY	1232	GCGAGGGGAGGCTAATCTCAGAAAGCGCGTGTGTGTCGCGAT	1273	
Db	42	GCGAGGGTGAAGCTAATCTCAAAAACCGTCTCTCAGTTCCGAT	1	



genomic survey sequence.  
ACCESSION BZ507023  
VERSION BZ507023.1 GI:27028590  
KEYWORDS  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 824)  
Ayele,M., Haas,B.J., Kumar,N., Mu,H., Xiao,Y., Van Aken,S.,  
Utecherback,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
JOURNAL Genome Res. 15 (4), 487-495 (2005)  
PUBMED 15805490  
COMMENT Other\_GSSs: BONIT76TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.  
FEATURES  
location/Qualifiers  
source 1..824  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO100DH3"  
/db\_xref="taxon:3712"  
/clone\_1lb="BO\_1.6 2 KB tot"  
/note="Vector: PHOS1; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into PHOS1 using BstXI linkers"  
ORIGIN  
Query Match 35.2%; Score 512.8; DB 9; Length 824;  
Best Local Similarity 78.1%; Pred. No. 8.7e-147;  
Matches 642; Conservative 0; Mismatches 177; Indels 3; Gaps 2;  
QY 484 GTGCCAGCAGCCCGGTAATACGTAAGGTCGAAGCGTTAATCGAATTAATCTGGCGCTAA 543  
DB 3 GTGCCAGCAGCCCGGTAATACGTAAGGTCGAAGCGTTAATCGAATTAATCTGGCGCTAA 62  
QY 544 GGGTCGCGAGCGCGCTTTGAAGTCAGATGTGAATCCCGGGCTTAACTCGGAATTCG 603  
DB 63 GCGTCTGTAGTGGCTTTTAAAGTCCCGCGTCAATCCAGGGCTCAACCTCGAGCAGGC 122  
QY 604 GTTTGAAACTCAAGGCTAGAGTGTGCAAGGAGGAGGTAATTCATGTGTAGCACTGA 663  
DB 123 GTTGAAGAACTCAAGGCTTAGTGTGAGTGGGAGGAGGAAATTTCCGGTGGAGCGGTGA 182  
QY 664 AATGCTAGAGATATGGAAGAACATCGATGCGGAGCAGCGCTTCGCGTTAATCACTGAC 723  
DB 183 AATGCTAGAGATATGGAAGAACACCAAGCGGAAACACACTCTCGTGGCGGACACTGAC 242  
QY 724 GCTTCATCAGAAAGCGTGGGAGCAAAACAGGATTAATCACTCTGTTAGTTCAGCGCCTA 783  
DB 243 ACTGAGAGACGAAGCTTAGGGGAGCGAATGGAATTAATCCCAAGTATGTCCTAGCCGTA 302  
QY 784 AACGATGCAACTGTGTTGGGGCTTATTAAGCGCTTG--GTAAAGAACTAAACGCTGAA 841  
DB 303 AACGATGATATAGGCGCTGTGGTATCGACCCGTGAGTGTCTGTACTTAACCGCTTAA 362  
QY 842 GTTGAACGCTTGAGGAGATACCGTGCAGATTAATACTCAAGGAATTAAGCGGAGACCG 901  
DB 363 GTATCCCGCTGGGAGATACGTTGCAAGATGAATCAAGGAATTAAGCGGAGACCG 422  
QY 902 CACAAGCGGTGATTAATGTGATTAATTCAGTCAAGCGGAAACCTTACTTACTCTTG 961  
DB 423 CACAAGCGGTGAGATGTGTTAATTCAGTCAAGCGGAAACCTTACTTACTTCTTG 482

QY 962 ACATGATGCAATTTTCTAGATATGATTAAGT-CTTCGGGAAACGCTAAACAGGTGCTG 1020  
DB 483 ACATCCCGGAATCTCTTGAAGAGAGAGGCTGCTTGGGAACCGGACACAGGTGGTG 542  
QY 1021 CATGCTGTCTGACAGCTCGTGTGAGATGTTGGGTTAAGTCCGCAACGAGCCCAAC 1080  
DB 543 CATGCTGTCTGACAGCTCGTGTGAGATGTTGGGTTAAGTCCGCAACGAGCCCAAC 602  
QY 1081 CTGTCAATTAATGTCATCAATTTGTTGGGACCTTTAATGACATCGCGGTGACAAACCG 1140  
DB 603 CTGTGTTAGTTGACACCGTTGAGTTGGAACCTTGAACACATCTCGGTGATTAAGCG 662  
QY 1141 GAGGAAGTGGGAGAGAGATGCAAGTCCATGAGCCCTTATGGGTGAGGCTTCAACGTAA 1200  
DB 663 GAGGAAGTGAAGATGACGTCAAGTCAATCAATGCCCCCTTATGCTCGGCGGACACACGTGC 722  
QY 1201 TACAATGGCGCTACAGAGGTTGCCAACCCCGAGGAGGAGCTAATCTCAGAAAAGCGCG 1260  
DB 723 TACAATGGCGGAGACAAAGGTCGCGATCCGCGAGGAGCTAACTCCAAACCCGT 782  
QY 1261 TCGTAGTCCGATCGAGAGTTCGAACCTGACTCCGTGAATGC 1302  
DB 783 CCTCAGTTGCAATGCAAGGCTGCACTCGCTGCATGAAGCC 824  
RESULT 63  
BH529859 837 bp DNA linear GSS 14-DEC-2001  
LOCUS BH529859  
DEFINITION BOGOD80TR BOGO Brassica oleracea genomic clone BOGOD80, genomic  
survey sequence.  
ACCESSION BH529859  
VERSION BH529859.1 GI:17745657  
KEYWORDS  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 837)  
Ayele,M., Haas,B.J., Kumar,N., Mu,H., Xiao,Y., Van Aken,S.,  
Utecherback,T.R., Mortman,J.R., White,O.R. and Town,C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
JOURNAL Genome Res. 15 (4), 487-495 (2005)  
PUBMED 15805490  
COMMENT Other\_GSSs: BOGOD80TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.  
FEATURES  
location/Qualifiers  
source 1..837  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO100DH3"  
/db\_xref="taxon:3712"  
/clone\_1lb="BOGO"  
/note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared  
genomic DNA inserted into PHOS1 using BstXI linkers"  
ORIGIN  
Query Match 35.2%; Score 512.8; DB 9; Length 837;  
Best Local Similarity 78.2%; Pred. No. 8.7e-147;  
Matches 654; Conservative 0; Mismatches 177; Indels 5; Gaps 3;  
QY 535 GGGGTTAAAGGTCGCGAGCGGCTTTGTAAATGATGTAATATCCCGGCTTAACCT 594

Db	2	GGGCGTAAAGCGTCTGTAGGTGGCTTTTAACTCCGCGGTCAAATCCAGGGCTCAACCC	61
QY	595	GGGAATTGCGTTTGGAACTACAAAGCTAGAGTGGCAGAGGAGGTGGAATTCATGTG	654
Db	62	TGGAAGGCGGCGTGAABACTTACCAAGCTTGATGCGTGGGCAAGGGAATTTCCGGG	121
QY	655	TAGCAGTGAATGCGTAGAGATATGGAAGACATGATGGCAGACCTCTCGGGT	714
Db	122	GAGCGGTGAATGCGTAGAGATCGAAGAACCAACGCGCGAAAGCACTGTGCGGGCC	181
QY	715	AACAATGACGCTCAGTGCACGAAGGGTGGGGAGCAACAGATTGATACCTCGTAGTC	774
Db	182	GACACTGACACTGGAAGACGAAGAAGCTAGGGGGCGAATGGGATTGATATCCCAATAGTC	241
QY	775	CACGCGCTAAACGATGTCAACTAGTTGTGTGGGCTTATTTAGGCTTG--GTAAAGAGCTA	832
Db	242	CTAGCGGTGAACGATGGAATCTAGGCGCGTGTGCGATGAGACCGCGTGCAGTGTGTAGCTA	301
QY	833	ACGCGTGAAGTTGACCGCGCTGGGGAGTACGCTGCAGAGTTAAACTCAAGGAATTGAC	892
Db	302	ACGCGTTAAAGTATCCCGCTGGGGAGTACGTTTCGGAAGATAAACTCAAGGAATTGAC	361
QY	893	GGGGACCGGCACAAGCGGTGATTTATGTGGATTATTTGATGTCACGCGGAATTCCTAC	952
Db	362	GGGGGCGCGCACAAACGGGTGAGCATGTGTATTATTCATGTCAAAGCGAAGACTTAC	421
QY	953	CTACCGTTGACATGTAGCGAATTTTCTAGAGATAGATTAGTG--CTTCGGGACGCTAAC	1011
Db	422	CAGGCGTTGACATGCGCGGAAATCCTTTGAAGAGAGGGGTGCTTCGGGACGCGAC	481
QY	1012	CAGGTGTCATGATGCTGTGTGTCAGCTCGTGTGCTGATGATTTGGTTAACTCCGCAACG	1071
Db	482	CAGGTGTGACATGCGCTGTGTGTCAGCTCGTCCGTAAAGGTGTGTGGTTAACTCCCAACG	541
QY	1072	AGCGAACCCTTGTGCTAATTAATGGCATATTGGTTGGCACTTAAAGACCTGCGGAT	1131
Db	542	AGCGAACCCTGTGTTTATGTGGCCACGTTGATGTTGAACTTGAACGACGTCGCGT	601
QY	1132	GACAAACCGSAGGAAGTGGGATGACGTCAGTCTTCATGCGCTTATGTGGTAGGGCTT	1191
Db	602	GATAGCCGGAGGAAGTGAAGTACGTCAGTCAATGATCATATGCGCTTATGCGCTGGGGA	661
QY	1192	CACACGTATATACATATGCGCGCTGACAGAGGTTGCCAACCCCGSAGGGAGACTTAATCTCA	1251
Db	662	CACAGTGTCTACAAATGCGCGGACAAAGGTTGCGATCCCGSAGGGTAGACTTAATCTCA	721
QY	1252	GAAAGCGGTCGTATGTCGGGATCGGAGTCTGCAATTCGACTCCGGAAGTCCGATTCCT	1311
Db	722	AAACCCGTCCTCACTTCGGATTGCAAGCTGCAATCCGCTGCAAGAAAGCGGAATCGCT	781
QY	1312	AGTATTCGCGATTCAGCATGTC--GCGGTGAATACGTTCCCGGCTCTGTATACACAC	1365
Db	782	AGTATTCGCGGTCAAGCATACAGGCGGTGATTCGTTCCGCGGCTCTGTATACACAC	837
RESULT 64			
BZ435283			
LOCUS			
BOHNF62T BO.1.6.2_KB_tot Brassica oleracea genomic clone BOHNF62,			
DEFINITION			
BZ435283			
ACCESSION			
BZ435283.1 GI:2686341			
KEYWORDS			
SSS.			
SOURCE			
Brassica oleracea			
Brassicaceae			
Brassicaceae			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;			
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.			
1 (bases 1 to 862)			
REFERENCE			
AUTHORS			
Aylele,M., Hase,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,			
Utechtbeck,T.R., Wortman,J.R., White,O.R. and Town,C.D.			
TITLE			
Whole genome shotgun sequencing of Brassica oleracea and its			

JOURNAL  
 PUBMED  
 COMMENT

Application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 15805490  
 Other\_GSSs: BONHE62TR  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TP  
 Class: sheared ends.

FEATURES  
 Source  
 1..862  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BONHE62"  
 /clone\_1fb="BO\_1.6\_2\_KB\_tot"  
 /note="vector: PHOS1; Site 1: BstXI; 1.6-2 kb sheared  
 total DNA inserted into PHOS1 using BstXI linkers"

ORIGIN

Query Match 35.1%; Score 512; DB 9; Length 862;  
 Best Local Similarity 79.1%; Pred. No. 1.6e-146;  
 Matches 634; Conservative 0; Mismatches 165; Indels 3; Gaps 2

448 ATGACGGTATGACGAGAAGACACCGGCTAACTACGTGCCAGACGCCGGTAAATACGT 507  
 60 ATGACGGTATCGGGGAATTAAGCATCGGCTACTCTGTGCCAGACGCCGGTAAATACAG 119  
 508 AGGGTGCAAGCGTTAATGGAATTACTGGCGTTAAAGGCTGCAGCGCGCTTTGTAAGT 567  
 120 AGGATGCAAGCGTTATCCGGAATGATTTGGCGTTAAAGCGTCTGTAGTGGCTTTTAAAT 179  
 568 CAGATGTGAATCCCCCGGCTTAACTCTGGGAATTCGTTTGAAATCTACAAGCTTAGACTG 627  
 180 CCGCGCTAAATCCCAAGGCTCAACCCCTGCACAGCGCGGTGAAACTACCAAGCTTGAGTA 239  
 628 TGGCAGAGGAGGTGGAATTCATGTGAGCAGTGAATTCGCTAGAGATTATGGAAGACA 687  
 240 CGGTAGGGGCGAAGGGAATTTCCGTGGAGCGGTGAATTCGTAGAGATTCGGAAGGAACA 299  
 688 TCGATGGCGAAGGCAAGCTCTCTGGGTTAACACTGACGCTCATGACGAAAGCGTGGGAG 747  
 300 CCAACGGGGAAGCACTCTGCTGGGCCGACACTGACCTGAGAGACGAAAGCTAGGGAG 359  
 748 CAAACAGATTAGATACCTCTGTAGTCCACGCGCTTAAAGATGTCACTAGTTGTTGGGC 807  
 360 CGAATGGGATTAGATACCCCAGTAGTCTTACCGCTTAAAGATGATGATCTAGGCGCTGTGC 419  
 808 CTATATTAGCTTG--GTACGAAGCTAACGGGTGAAGTTGACCGGCTGGGGAGTACGGTC 865  
 420 GTATTCAGCCCGTGCAGTCTGTAGTAAAGGCTTAAGATCCCGCTGGGGAGTACGTTTC 479  
 866 GCAAGATTAAACTCAAGGAATTGACGGGAGACCGGCAAGCGGTGATTATGTGATT 925  
 480 GCAAGAAATGAATCTAAAGAAATTGACGGGGGCCGCGCAAGCGGTGAGCATGTGATT 539  
 926 AATTGATGACACGGGAAAACCTTACCTACCTTTGACATGTAGGAAATTTTCTAGAGAT 985  
 540 AATTGATGACAAAGGAACCTTACCTACCGGGCTTGAATATCCCGGAATTCCTTTGAAAG 599  
 986 AGATTATAGTG-CTTGGGGAACGCTAACAGGTGCGATGTGCTGCGACGCTGTGTCG 1044  
 600 AGAGGGGTGCTTGGGGAACGCGGACAGGTGTGCAATGTGCTGCTAGCTGCTGCG 659  
 1045 TGAGATGTTGGGTTAAGTCCCGCAACGAGCCCAACCTTGTCTAATTAATGGCATTAATTG 1104  
 660 TAAGGTGTTGGGTTAAGTCCCGCAACGAGCCCAACCTCTGTGTTAAGTTGGCAACGTTGA 719

QY 1105 GTTGGACCTTAATGAGCTGCGGTGACAAACCGAGAGAGTGGGATGACTCAAG 1164  
 DB 720 GTTTGGAACCTTGAAACAGACTGCGGTGATTAAGCCGAGAGAGTGAAGATGACTCAAG 779  
 QY 1165 TCTCATGGCCCTTATGAGGCTTCAACACTAATACAAATGCGCCGTACAGAGGCTTG 1224  
 DB 780 TCATATGCGCCCTTATGAGGCTTCAACACTAATGCGCCGTACAGAGGCTTG 839  
 QY 1225 CCAACCGGAGGAGGAGCTAA 1246  
 DB 840 CCAATCCGAGGAGGAGCTAA 861  
 RESULT 65  
 BZ463792 1030 bp DNA linear GSS 13-DEC-2002  
 LOCUS BZ463792  
 DEFINITION BONK105TR BO\_1.6\_2\_KB\_tot Brassica oleracea genomic clone BONK105,  
 genomic survey sequence.  
 ACCESSION BZ463792  
 VERSION BZ463792.1 GI:26749971  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Brasicaceae  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 1030)  
 Ayala, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
 Utecherback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 15805490  
 JOURNAL  
 PUBMED  
 COMMENT Other\_GSSs: BONK105TF  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.  
 location/Qualifiers  
 1..1030  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="T01000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BONK105"  
 /note="Vector: pHD51; Site 1: BexXI; 1.6-2 kb sheared  
 total DNA inserted into pHD51 using BexXI linkers"  
 ORIGIN  
 Query Match 35.1%; Score 512; DB 9; Length 1030;  
 Best Local Similarity 77.3%; Pred. No. 1.6e-146;  
 Matches 660; Conservative 0; Mismatches 190; Indels 4; Gaps 3;  
 QY 608 GAAACTACAGGCTAGAGTGTGAGAGGAGGAGTTCATGTGTAGACGTGAATG 667  
 DB 2 GAAACTACAGGCTAGAGTGTGAGAGGAGGAGTTCATGTGTAGACGTGAATG 61  
 QY 668 CATTAGATATGAAAGCAATCCATGCGAGGAGCACTCTGCGGTTAAACGTAGCGTC 727  
 DB 62 CGTAGAGATGCGAAAGAACACCAACGCGGAAAGCACTCTGCGGCGACACTGACTG 121  
 QY 728 ATGACGAAACCGTGGGAGCAACAGAGTTAGATACCTGTAGTCCACCCCTAAACG 787  
 DB 122 AGAATCGAAAGCTAGGAGGAGCAATGAGATTAGATACCCAGTAGTCTTACCCGTAACG 181  
 QY 788 ATGTCAACTAGTGTGGGCTTATTAAGGCTTG--GTAAAGAGCTAACGCGTGAAGTTG 845

DB 182 ATGATACTAGGCGGTGTCGATATGACCCGTCAGTGTAGCTAAACGCTTAAGTAT 241  
 QY 846 ACCGCTGGGAGATACGCTCGCAAGATTAAATCTAAAGAAATTGACGGGACCCGACA 905  
 DB 242 CCGCCTGGGAGATACGCTCGCAAGATTAAATCTAAAGAAATTGACGGGACCCGACA 301  
 QY 906 ACCGCTGGATTTATGAGATTTAATTGATGCAACGGGAAACCTTAACCTTACAT 965  
 DB 302 ACCGCTGGAGCATGTGTGTTAATTGATGCAACGGGAAACCTTAACCTTACAT 361  
 QY 966 GTAGCAATTTTCTAGATATGATGATG-CTTCGGGACGCTTAACACAGTGTCTGATG 1024  
 DB 362 GCCGCAATCTCTTGAAGAGAGGGGTCTTCGGGACGCGACACAGTGTCTGATG 421  
 QY 1025 GCTGTGTCAGCTCGTGTCTGATGATGTTGGGTTAAGTCCCGACACAGCCGCAACCTTG 1084  
 DB 422 GCTGTGTCAGCTCGTGTCTGATGATGTTGGGTTAAGTCCCGACACAGCCGCAACCTTG 481  
 QY 1085 TCATTAAATTGCCATCATTTGTTGGCACTTTAATGAGACTGCGGAGCAAAACCGAGG 1144  
 DB 482 TGTTAAGTTGCCACGTTGATGTTGAAACCTGACAGACTGCGGAGTAAAGCCGAGG 541  
 QY 1145 AAGTGGGATGACGTCAGTCCCTCATGSCCTTATGAGGCTTCAACGTAATTACA 1204  
 DB 542 AAGTGGGATGACGTCAGTCCCTCATGSCCTTATGAGGCTTCAACGTAATTACA 601  
 QY 1205 ATGGCGCGTACAGAGGCTTCCAAACCGCGAGGGGAGCTAATCTCAGAAAGCGCTCGT 1264  
 DB 602 ATGGCGCGGACAAAGGCTCGCGATCCCGAGAGGCTAGCTAATCTCAGAAAGCGCTCGT 661  
 QY 1265 AGTCCGAGATCGAGTGTGCAACCTCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 1324  
 DB 662 AGTTCGATATGCAAGGCTGCAACTGCTGTGATGAAACCGGATTCGTAATTCGCGCT 721  
 QY 1325 CAG-CATGTCGCGGTGAATACGTTCCCGGCTTGTGACACACCCGCTGACCATATG 1383  
 DB 722 CAGCATACGCGCGGTGAATACGTTCCCGGCTTGTGACACACCCGCTGACCATATG 781  
 QY 1384 AGTGGGTTTCAACAGAGCAGTGTGTAACCTTAAGAGAGGCGCTTGCACCGTGAAT 1443  
 DB 782 AGCTGGCAGATGCCGAAAGTGTGTAACCTTAAGAGAGGCGGCTGCGAAGGAGGCG 841  
 QY 1444 TCATGACTGGGCTG 1457  
 DB 842 TAGTGACTGGAATG 855  
 RESULT 66  
 BZ459110 1082 bp DNA linear GSS 13-DEC-2002  
 LOCUS BZ459110  
 DEFINITION BONBR59TR BO\_1.6\_2\_KB\_tot Brassica oleracea genomic clone BONBR59,  
 genomic survey sequence.  
 ACCESSION BZ459110  
 VERSION BZ459110.1 GI:26740285  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Brasicaceae  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 1082)  
 Ayala, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,  
 Utecherback, T.R., Wortman, J.R., White, O.R. and Town, C.D.  
 Whole genome shotgun sequencing of Brassica oleracea and its  
 application to gene discovery and annotation in Arabidopsis  
 Genome Res. 15 (4), 487-495 (2005)  
 15805490  
 JOURNAL  
 PUBMED  
 COMMENT Other\_GSSs: BONBR59TF  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208

Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR  
Class: sheared ends.

## FEATURES

source

Location/Qualifiers  
1..1082

/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONBR59"  
/clone\_1lb="BO\_1.6-2\_KB\_tot"  
/note="Vector: pHO51; Site 1: BexXI; 1.6-2 kb sheared  
total DNA inserted into pHO51 using BexXI linkers"

## ORIGIN

```

Query Match      35.1%; Score 512; DB 9; Length 1082;
Best Local Similarity 73.6%; Pred. No. 1.7e-146;
Matches 742; Conservative 0; Mismatches 215; Indels 51; Gaps 5;

3  TGAACGCTGCGGCATGCTTTTACACATGCAATGCAACGGCAGCAGCATGCTGCATCT 62
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
124 TGAACGCTGCGGCATGCTTTTACACATGCAATGCAACGGCAGCAGCATGCTGCATCT 171
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63  GGTGGCAGTGGCGGACGCGGTGATGATGATCGAATCGAATCGATTCAGAAAGAGGGGGTATC 122
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
172 -GTTTCACTGGCGGACGCGGTGATGATGATCGAATCGAATCGATTCAGAAAGAGGGGGTATC 230
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
123 GCATGAAAGATGTCTTAATCCCATATCTCTTAAGAGAAAGCAGGGGATCGAAGA 182
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
231 AGCTGAAACGGCTGCTTAATACCCC -GTAGCTGAGAGCAAAAGAGGAATC----- 282
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
183 CTTGGCGCTTTGGAGCGCCCATGCTCTATTAAGCTAAGTTGGTGGGGTAAAGCCCTACCA 242
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
283 -----CGCCGAGAGGGGCTCGCGCTTATTAGCTGATGTTGGTGAAGCAATAGCTTACCA 337
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
243 AGCGCAGATCATGATGTTGTTGAGAGAGCAGCAGCACTGGGATCTGAGACAGCGC 302
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
338 AGGCGATGATCATGATGTTGTTGAGAGAGCAGCAGCACTGGGATCTGAGACAGCGC 397
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
303 CCAGACTCTCTTACGAGAGGAGCAGAGTGGGAAATTTTGGACAAATGGGCGAAGCTGATCCA 362
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
398 CCAGACTCTCTTACGAGAGGAGCAGAGTGGGAAATTTTGGACAAATGGGCGAAGCTGATCCA 457
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
363 GCAATGCCGCGTGAAGAAAGCCCTTGGGTTGAAGCTCTTCACTCAGAAAGAA 422
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
458 GCAATGCCGCGTGAAGAAAGCCCTTGGGTTGAAGCTCTTCACTCAGAAAGAA-- 515
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
423 AGGTTACGGTAATATCTGACTCATGACGGTATCGAAGAAAGCAGCGCTTAATCTA 482
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
516 -----AGCAATGACGGTATCTGGGAAATTAAGCATCGGCTTAATCT 554
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
483 CGTGCAGAGCGCGGTAAATCGTAGGGTGAAGGCTTAATCGAATTAATCTGGGCGTAA 542
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
555 TGTGCAGAGCGCGGTAAATCGTAGGGTGAAGGCTTAATCGAATTAATCTGGGCGTAA 614
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
543 AGGTCGCGAGCGGCTTTGTAAGTCAAGTGAAGTCCCGGGCTTAACTCGGGAATTG 602
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
615 AGCGCTGATGAGTCTTTTAAAGTCCGCGTCAATTCAGAGGCTCAACCTCGAGCAGG 674
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
603 CGTTTGAATCTCAAGGCTAGAGTGGCAGAGGAGGTGAATTCAATGTGTAGCAGTG 662
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
675 CGGTGAAATCTCAAGGCTAGAGTGGCAGAGGAGGTGAATTCAATGTGTAGCAGTG 734
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
663 AAATCGTAGAGATGGAAGAAATCGTAGGGCAGAAAGCGCTCTCGGGTTAATCACTGA 722
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
725 AAATCGTAGAGATGGAAGAAATCGTAGGGCAGAAAGCGCTCTCGGGTTAATCACTGA 794
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
723 CGCTATGACGAAAGCGTGGGAGCAACAGATTTAATACCTCGTGTAGTCAAGCCCT 782
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
795 CACTAGAGAGCAAAAGCTAGGGGAGCGAATGGGAATTAATCCCAAGTAGTCTTAGCGGT 854
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
783 AAACATGTCATAGTTGTTGGGCTTATTAGGCTTG--GTAGCAAGCTAAACGCTGA 840

```

```

Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
855 AAACATGTCATAGTTGTTGGGCTTATTAGGCTTG--GTAGCAAGCTAAACGCTGA 914
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
841 AGTTACCGGCTGGGAGTACGCTGCGAAGATTAATACTCAAGAGATTTAGCGGGAGCC 900
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
915 AGTATCCCGCCCTGGGAGTACGCTTGCAGAGATGAATCAAGAGATTTAGCGGGAGCC 974
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
901 GCAGAGCGGTGATATGATGATTAATTCGATGCAAGCGGAAACCTTACCTTACCTT 960
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
975 GCAGAGCGGTGATGATGATTAATTCGATGCAAGCGGAAACCTTACCTTACCTT 1034
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
961 GACATGACGGAATTTTCTAGAGATGATTAAGTGTCTGGGAACGCTA 1008
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1035 GACATGACGCGGAATCTCTTGAAGAGAGGGGTGCTTGGGAACGCGA 1082

```

RESULT 67  
BH484054  
LOCUS  
DEFINITION  
BOHGA75TR BOHG Brassica oleracea genomic clone BOHGA75, genomic survey sequence.  
ACCESSION  
BH484054  
VERSION  
BH484054.1 GI:17692158  
KEYWORDS  
SOURCE  
ORGANISM  
Brassica oleracea  
Brassica oleracea  
Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, eudicotyledons, core eudicotyledons, rosids, eurosids II, Brassicales, Brassicaceae, Brassica.  
1 (bases 1 to 819)  
Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utecher, T.R., Mortman, J.R., White, O.R. and Town, C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
Other GSSE: BOHGA75TF  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

## FEATURES

source

Location/Qualifiers  
1..819  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOHGA75"  
/clone\_1lb="BOHG"  
/note="Vector: pHO51; Site 1: BexXI; 2-3 kb sheared  
genomic DNA inserted into pHO51 using BexXI linkers"

## ORIGIN

```

Query Match      35.1%; Score 511.4; DB 9; Length 819;
Best Local Similarity 78.1%; Pred. No. 2.4e-146;
Matches 640; Conservative 0; Mismatches 176; Indels 3; Gaps 2;

512 TGAACGCTTAAATCTGAGGCTTAAGGTTGGCAGGGGCTTTGAAGTCA 571
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1  TGAACGCTTAAATCTGAGGCTTAAGGCTTGAAGGCTTTTAAAGTCCG 60
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
572 TGTGAATCCCCGGGCTTAACCTGGGAATTCGTTTGAATCAAGGCTAGAGTGTGC 631
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61  CGTCAATCCCAAGGCTCAACCTGGAACGCGGTGAAATCAACAGCTTAGTAGCGT 120
Qy | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
632 AGAGGAGGTGAATTCATGTGACAGTGAATTCGCTGAGATGAGATGAGAAACATCGA 691
Db  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 AGGGCAGAGGGAATTTCCGTGAGCGGTGAATGCGTAGAGATCGGAAGAAACACCA 180

```

Oy		692	TGGGGAAGGCACGCTCCTCGGGTTTAAACA	CTAAGCTCATGTGACAGAAACCTGTGGGAGCAAA	751
Db		181	CGGGAAAAGCACTTGCTGGGCCCAACTGACAT	CACTGAGAACGAAACCTTAGGGGACGGA	240
Oy		752	CAGATTAGTAATACCCTGGTAGTCACAGCC	CTTAAACGATGTCAACTAGTTGTCGCTTA	811
Db		241	TGGGATTATAGTAACCCACAGTAGTCTT	AGACCGGTAACGATAGTAATAAGGCGCTGTG	300
Oy		812	TTAGGCTTG--GTAAACGAGCTTAACCGGT	GTAAGTTGACCCGCTGGGAGATACCGGTG	869
Db		301	CGACCCGGCACGTGCTGTAGCTTAACGG	CTTAAAGTATCCCCTGTGGGAGATACGTT	360
Oy		870	GATTAAAACTCAAAGAATTGACGGGAGC	CCGGCAACAGCGGTGATTAATGTGATTAAT	929
Db		361	GAATGAATCTCAAAAGAAATTBACGGGG	CCCCGGCACAGCGGTGAGATGTGTTAATT	420
Oy		930	CGATGCAACGGAAAAACCTTACCTTAC	CTTACATGTAGCGAAATTTTCTAGATATGAT	989
Db		421	CGATGCAAAAGGAAAGAACCTTACACG	AGGCTTGAATGATCCGCTTGAAGAGAG	480
Oy		990	TATGTG-CTTCCGGAAACGTTAACACAG	GTCTGTGACATGCTGTGCTGTGCTGTGAG	1048
Db		481	GGGTGGCTTCCGGGAACGGGACACAGG	TGTGTGATGCTGTGTGATGCTGTGCGTGA	540
Oy		1049	ATGTTGGGTTAAGTCCCGCAACGAGCG	CAACCCCTTGTCACTAATATGTCANTATTTG	1108
Db		601	GGAACCTTGAAACAGACTCCCGGTGAT	AAGCCGAGGAAGTGAAGATGACTCAAGTCA	660
Oy		1169	CATGCCCCCTTATGSGTAGGGCTTCA	CACGTAAATCAATGGGCGCTACAGAGGTTG	1228
Db		661	CATGCCCCCTTATGCCCCCTTAGCGAC	ACACGTGTCAAAATGGCCGGGACAAAGGTC	720
Oy		1229	CCCCGGAGGGGAGCTTAATCTCAGAA	ACGGGTCGTAGTCCGSATCGSAGTCTGCA	1288
Db		721	CCCCGGAGGGTACGCTTAATCCA	AAAACCCGTCCTCAATTTGGATGTGACG	780
Oy		1289	GACTCCGTGAAGCTCGAAATCGCTA	GTAAATCGCGATCAG 1327	
Db		781	GCCTGCATGAAGCCGGAATCGCTA	GTAAATCGCGCTCAG 819	
RESULT 68		BH582045	861 bp	DNA	linear
BH582045/c		BOGZU46TF	BOGZ Brassica oleracea genomic clone BOGZU46,		genomic survey sequence.
LOCUS		BH582045			
DEFINITION		BH582045.1	GI:17834502		
ACCESSION		BH582045			
VERSION		BH582045.1			
KEYWORDS		GSS.			
SOURCE		Brassica oleracea			
ORGANISM		Brassica oleracea			
REFERENCE		1 (baaes 1 to 861)			
AUTHORS		Ayele,M., Haae,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Utecherback,T.R., Wortman,J.R., White,O.R. and Town,C.D.			
TITLE		Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res. 15 (4), 487-495 (2005)			
JOURNAL		PUBMED			
COMMENT		Other GSSs: BOGZU46TR			
		Contact: Chris Town			
		TIGR			
		9712 Medical Center Drive, Rockville, MD 20850, USA.			
		Tel: 301-838-3523			
		Fax: 301-838-0208			
		Email: cdtown@tigr.org			

**FEATURES**  
**SOURCE**

DNA is from a doubled haploid provided by Tom Osborn  
Seq primer: TF  
Class: sheared ends.

Location/Qualifiers  
1. .861

```
/note="Vector: pHSO1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHSO1 using BstXI linkers"
```

## ORIGIN

Query Match	35.1%;	Score 510.8;	DB 9;	Length 861;
Best Local Similarity	76.5%;	Pred. No. 3.7e-146;		
Matches 676;	Conservative 0;	Mismatches 182;	Indels 26;	Gaps 3

OY	243	AGGCGACGATCAGTAGTGTGTTCTGAGAGGACGAC	CACGCCACACTGGGACCTGAGACACGGC	302
Db	861	AGCGGATGATCAGTAGCTGCTCCGAGAGGATGAT	CACGCCACACTGGGACCTGAGACACGGC	802
OY	303	CCAGACTCTTACCGGAGAGCAGCTGGGGAAATTT	TGACAATGGGGCGCAAGCCTGATCCA	362
Db	801	CCAGACTCTTACCGGAGAGCAGCTGGGGAAATTT	CCGCAATGGGGCGCAAGCCTGACGGA	742
OY	363	GCAATGCGCGGTGAGTAGAAGGCTTCGGGTTGT	AAAGCTCTTTCACTGTCGAGAGAAA	422
Db	741	GCAATGCGCGGTGAGTAGAAGGCTTCGAGGCTTA	CGGATCTTTCCAGAGAGAA--	684
OY	423	AGTTTACGGTAAATTAATCTGTA	CTCATGACGGTATGCAAGAGAGACACCGGCTAACTA	482
Db	683	-----	AGCAATGACGGTATCTGGGGAAATTAACATCTGGCTAACTC	645
OY	483	CGTGCCAGACGACCCCGGTAATACGTAGGGTCCA	AGCGTTAATCGGAATTCCTGGGCGTTAA	542
Db	644	TGTGCGACAGACGCCCGTAATACAGAGATGCA	AGCGTTAATCCGGAATGATTTGGCGTTAA	585
OY	543	AGGGGCGCAGCGCGCTTTGTAACTCAGATGTGA	AAATCCCGGCTTAACTCTGGGAAATYG	602
Db	584	AGCGTCTGTAGGTGGCTTTTAACTCCGCGCT	CAAAATCCAGGGCTCAACCTGTGACAGG	525
OY	603	CGTTTGAATCTACAAGGCTAGAGTGTGGCAG	AGGAGGTGAATTCATGTGTGACAGTG	662
Db	524	CGGTGGAACCTACCAAGCTTAGTACGGTAGGG	CAGAGGGAATTTCCGCGTGGAGCGGTG	465
OY	663	AAATGCGTAGAGATATGGAAGAAACATCGAT	GTGGCGAAAGGCCCTCTGGGTTACACTGA	722
Db	464	AAATGCGTAGAGATATGGAAGAAACCAACGCG	CAAAAGCACTGTGTGGCGGACACTGA	405
OY	723	CGCTCATGACGAAAGCGTGGGGAGCAAA	CAGAAATTAGATACCCGTGATGCAACGCCCT	782
Db	404	CACCTAGAGAGCAAAAGCTTAGGGGAGCAAA	TGGGATTAATACCCAGTAGTCTTAGCGGT	345
OY	783	AAACGATGTCAACTAGTGTGGGCTTTATTAG	CGCTTG--GTAAAGAAAGCTTACGCGTGA	840
Db	344	AAACGATGATTAATGAGCGCTGTGGGTAT	CGAACCCGTGCTGTAGCTTACGCGCTTA	285
OY	841	AGTTGACCGCCTGGGGAGTACGGTTCGCAAG	TTAAACTCAAAGGAATTTGACGGGGACCC	900
Db	284	AGTATCCCGCCTGGGGAGTACGGTTCGCAAG	ATGAACTCAAAGGAATTTGACGGGGGCC	225
OY	901	GCACAAAGGGGTGATTTATGTGAATTAAT	TCATGCAACGCGCAAAAACCTTACTCACTT	960
Db	224	GCACAAAGGGGTGAGACTGTGTTTAATTCAT	GCAAAAGCAAGAACTTACACAGGGCTT	165
OY	961	GACATGTAGCGAATTTTCTAGAGATGATTA	GTGTG--CTTGGGAAACGCTTAACACAGGTCT	1019
Db	164	GACATGTAGCGGAAATCTCTTGAAGAAAG	AGGGGTGCTTTCGGGAAACGCGGACACAGTGTGT	105
OY	1020	GCATGGCTGTCTCAGCTGTGTCTGTGAT	TTGGTTAAGTCCCGCAACGAGGCGCAAC	1075

Db 104 GCATGCTGTGCTGACGCTGTCGCCGTAAGTGTGGTTAAGTCCCGCAAGAGCGCAAC 45  
Qy 1080 CCTGTTCATTAATTGCCATCATTTGGTTGGCACTTTAATGAGA 1123  
Db 44 CCTCGTTTAAAGTGGCCACCGAAGATTGGAAACCTTGAAACAGA 1

RESULT 69  
B2468986 935 bp DNA linear GSS 13-DEC-2002  
LOCUS BOMKD78TF BO.1.6.2 KB tot Brassica oleracea genomic clone BOMKD78,  
DEFINITION genomic survey sequence.  
ACCESSION B2468986  
VERSION B2468986.1 GI:26764518  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 935)  
Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Utecherback,T.R., Wortman,J.R., White,O.R. and Town,C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
13805490  
Other GSSs: BOMKD78TR  
CONTACT: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source location/Qualifiers  
1..935  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOMKD78"  
/clone\_1lb="BO\_1.6.2\_KB\_tot"  
/note="vector: PHOS1; Site 1: BactI; 1.6-2 kb sheared  
total DNA inserted into PHOS1 using BactI linkers"

ORIGIN  
Query Match 35.0%; Score 509.4; DB 9; Length 935;  
Best Local Similarity 77.5%; Pred. No. 16-145;  
Matches 655; Conservative 0; Mismatches 186; Indels 4; Gaps 3;

Qy 617 AGCTAGAGTGTGGCAGAGGAGGAGTGGAAATTCATGTGTGAGAGTGAATGCGTAGAGT 676  
Db 1 AAGCTTGTAGTACGTGAGGAGGAGGAGGAAATTCGCTGAGAGGCGTGAATGCGTAGAGT 60  
Qy 677 ATGAAGAATCATGATGGGAGGAGGAGCGCTCTGGGTTAAACATGACGTGACGAGAA 736  
Db 61 CGGAAGAACAACCAACGCGGAGAAAGCACTCTGCTGGGCGACACTGACACTGAGAGAGAA 120  
Qy 737 AGCGTGGGAGGAGAAACAGATTAAGTATCCCTGTGTAGTCCAGCCCTTAAACGATGCACT 796  
Db 121 AGCTAGGGGAGGAGGAGGAGGAGTATGATCCAGTAGTCTTACCGGTAAACGATGATCT 180  
Qy 797 AGTTGTGGGCTTATTAGGCTTG--GTAAAGAACTAAACGCGTGAAGTTGACCGCTGG 854  
Db 181 AGGCGCTGTGCTATCGACCTCGACAGTCTGTACTTAACGCGTTAAGATATCCCGCTGG 240  
Qy 855 GGAGTACGCTGCGCAAGATTAAACTCAAGGAATTGACGCGGAGCCCGCAACAGCGGTGGA 914  
Db 241 GGAGTACGCTGCGCAAGATTAAACTCAAGGAATTGACGCGGAGCCCGCAACAGCGGTGGA 300

Qy 915 TTATGTGATTAATTCGATGCAACGCGAAAAAACCCTTACCTTACCTTGACATGAGCGAAT 974  
Db 301 GCATGTGTTTATATTCATGCAAGCAAGAAACCTTACAGAGGCTTGAATGCCGCGAAT 360  
Qy 975 TTTCTAGATATGATTAATG-CTTCGAGAACGCTTAACAGAGTCTGATGCTGTCTC 1033  
Db 361 CTTCTGAAAAGAGAGGGGTGCTTCGGGAGCGCGAGACAGAGTGTGATGCTGTCTC 420  
Qy 1034 AGCTGTGTGCTGAGATTTGGTTAACTCCCGCAAGAGGCGAAACCTTCTCATTAAT 1093  
Db 421 AGCTGTGTGCTGAGATTTGGTTAACTCCCGCAAGAGGCGAAACCTTCTGTTAAT 480  
Qy 1094 GCCATCATTTGTGTTGGCACTTTATATGACTGCGCGTGAACAACCGAGAGAGGTGGGG 1153  
Db 481 GCCACCTTGTGATTTGGAACTCTGAACAGACTGCGCGTGAATAGCCGAGAGAGTGAAG 540  
Qy 1154 ATGACGTCAAGTCTCATAGGCGCTTATGAGTGAAGGCTTACAGCTAATATCAATGCGCGT 1213  
Db 541 ATGACGTCAAGTCAATCATAGGCGCTTATGAGTGAAGGCTTACAGCTAATATCAATGCGCGG 600  
Qy 1214 ACAAGAGGTTGCCAACCGCGAGGGGAGCTAATCTCAGAAAGGCGCTGTATGCTCCGAT 1273  
Db 601 ACAAGAGGTTGCCAATCCCGAGGGGAGCTAATCTCAGAAAGGCGCTGTATGCTCCGAT 660  
Qy 1274 CGGAGTGTCAACTGCACTCCGTAAGTCCGTAAGTCCGTAAGTCCGTAAGTCCGTAAGTCCGTAAG 1332  
Db 661 TCGAGGCTGCACTGCTGCTGATGAGCGGAAATCGTATGATATGCGCGGTCAAGCAATAC 720  
Qy 1333 CGCGGTGAATACGTTCCCGGCTTTGTACACACCGCGCTTACACATGAGGAGTGGGTTT 1392  
Db 721 GCGGTTGAATTCGTTCCCGGCTTTGTACACACCGCGCTTACACATGAGGAGTGGGTTT 780  
Qy 1393 CACCAAGAGCGATGATTAACGTAAGAGGAGGCTTCCACGCTGATGATTCATGACTG 1452  
Db 781 TGCCGGAAGTGTTAACCTTAACGCAAGAGGAGGAGTGTGCGAAGCAAGGCTAGTACTG 840  
Qy 1453 GGGTG 1457  
Db 841 GAGTG 845

RESULT 70  
BH597210 861 bp DNA linear GSS 15-DEC-2001  
LOCUS BOHEB62TR BOHE Brassica oleracea genomic clone BOHEB62, genomic  
DEFINITION survey sequence.  
ACCESSION BH597210  
VERSION BH597210.1 GI:17849662  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 861)  
Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,  
Utecherback,T.R., Wortman,J.R., White,O.R. and Town,C.D.  
Whole genome shotgun sequencing of Brassica oleracea and its  
application to gene discovery and annotation in Arabidopsis  
Genome Res. 15 (4), 487-495 (2005)  
15805490  
Other GSSs: BOHEB62TF  
CONTACT: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.  
location/Qualifiers  
1..861

```
/organism="Brassica oleracea"  
/mol_type="genomic DNA"  
/strain="TO100DH3"  
/db_xref="taxon:3712"  
/clone="BOHER62"  
/note="Vector: PHOS1, Site_1: BatX1, 2-3 kb sheared  
genomic DNA inserted into PHOS1 using BatX1 linkers"
```

## ORIGIN

```
Query Match      34.9%; Score 508.8; DB 9; Length 861;  
Best Local Similarity 77.6%; Pred. No. 1.5e-145;  
Matches 666; Conservative 0; Mismatches 187; Indels 5; Gaps 4;  
  
QY 554 GGGGCTTTGTAAGTCAGATGTGAATCCCGGCGCTTAACCTGCGGAATTGCGTTGAAACT 613  
DB 3 GTGGCTTTTAAATCGCCGCTCAAAATCCAGGGCTCAACCTGGAACAGCGCGTGAACACT 62  
QY 614 ACNAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTAGCAGTGAATGCGTAGA 673  
DB 63 ACCAAGCTTAGAGTACGGTAGGGCAGAGGGAATTCGCGTAGAGCGGTGAATGCGTAGA 122  
QY 674 GATATGAAAGAACATGATGAGCCGGAAGGCAAGCTCTTGGGTTAACTGAGGCTCATGCAC 733  
DB 123 GATCGGAAAGAACACCAACCGCGAAGCACTGTGCGGCCGACACTGACACTGAGAGAC 182  
QY 734 GAAAGGTGGGGAGCAACAGATTAGATACCTGTGATGCAAGCCCTAAAGATGCA 793  
DB 183 GAAAGCTAGGGGAGCGAATGGAAATTAGATACCCAGTAGTCTAGCCGTAAAGATGGAT 242  
QY 794 ACTAGTTGTGGGCTTATTAGGCTTG--GTAAAGAGCTAACGCGTGAAGTTGACCGCC 851  
DB 243 ACTAGGCGCTGTGGTATGACCCGCTGACAGTGTGATGCTAACGCGTTAGTATCCGCC 302  
QY 852 TGGGAGTACGGTGCAGAAATTAAACTCAAGGAATTGACGGGACCCGCAAGGGGT 911  
DB 303 TGGGAGTACGGTGCAGAAATGAACTCAAGGAATTGACGGGACCCGCAAGGGGT 362  
QY 912 GGAATTATGGATTAATTGATGCAACGCGAAAACTTACTACCTTGACATGTAGCG 971  
DB 363 GGAACATGTGTTAATTCATGCAAGCGAAGAACTTACAGGGCTTGACATGCGCGG 422  
QY 972 AATTTCTAGAGATAGATTAGTG-CTTCGGGAACGCTAACACAGGTGCTGATGGTGC 1030  
DB 423 AATCTCTTGAAGAGAGGGGTGCTTCGGGAACGCGAACACAGGTGCTGATGGTGC 482  
QY 1031 GTCAGCTGTGTGCTGATGATGTTGGTTAAGTCCCGCAAGAGCGCAACCTTGTCTTA 1090  
DB 483 GTCAGCTGTGTGCTGATGATGTTGGTTAAGTCCCGCAAGAGCGCAACCTTGTCTTA 542  
QY 1091 ATTGCCATCAATTGGTTGGGCACTTATATGACTGCGCGGTGAACAAACCGAGGAGGTG 1150  
DB 543 GTTGCACCGGTGATTGGAACCTTGAACGACTGCGGGTGAATAGCCGAGGAGAGGTG 602  
QY 1151 GGGATGACGTCAAGTCTCAATGAGCCCTTATGGGTAGGGCTTCAACGTTAATACATGGCG 1210  
DB 603 AGGATGACGTCAATGATCAATGAGCCCTTATGGCCCTTGGGCGACACAGTGTCAATGGCC 662  
QY 1211 CGTACAGAGGGTTCCAAACCGCGAGGGGAGCTAATCTCAGAAAAGCGCTGTAGTCCG 1270  
DB 663 GGGACAAAGGGTTCGCAATCCCGCGAGGGGTGAGCTAACCAAAAACCGTCTCAATTCG 722  
QY 1271 GATCGAGTCTGCAACTCGACTCCGTAAGTCGGAATCGCTAGTAATGCGGATCAG-CA 1329  
DB 723 GATTGCAAGGCTGTCACTGCTGTGATGAAGCCGGAATGTGATGATGCGCGTCAAGCCA 782  
QY 1330 TGTCCGGTGAATAGTCCCGGGTCTTGTACACAGCCGCCGTCAACCATGAGAGTGGG 1389  
DB 783 TACGCCGCTGAATTCGTTCCCGGGGCTGTACAC-CCGCCGTCACTATGGGAGCTGG 841  
QY 1390 TTTCACCAAGACAGGTA 1407  
DB 842 CCATGCCCGAAGTCTGTTA 859
```

Search completed: May 5, 2006, 17:00:25  
Job time : 6226 secs

**THIS PAGE LEFT BLANK**









TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 612-339-0331  
 TELEFAX: 612-339-0361  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1536 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: rRNA  
 ORIGINAL SOURCE:  
 ORGANISM: Pseudomonas testosteroni  
 US-08-114-695A-7

Query Match 76.7%; Score 1118.2; DB 2; Length 1536;  
 Best Local Similarity 70.0%; Pred. No. 0;  
 Matches 1022; Conservative 252; Mismatches 178; Indels 9; Gaps 4;

1 ATTGAACGCTGGCGGCGAT-GCTTTACACATGCAAGTCCGAACGGCAGCAAGGATGCTTGA 59  
 28 AATGACGCTGGCGGCGATGCTTTACACATGCAAGTCCGAACGGCAGCAAGGATGCTTGA 84  
 60 TCTGGGCGGAGTGGCGGCGGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 119  
 85 -AATGACGCTGGCGGCGGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 143  
 120 AACGATCGAAGATGCTGCTAATACCGCATATCTTAAAGAGAAAGCAAGGAGATCGAA 179  
 144 AACGATCGAAGATGCTGCTAATACCGCATATCTTAAAGAGAAAGCAAGGAGATCGAA 203  
 180 AGACCTTGGCTTTTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 239  
 204 GGGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 263  
 240 CCAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299  
 264 CCAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 323  
 300 GGGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 359  
 324 GGGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 383  
 360 CCAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419  
 384 CCAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443  
 420 AAAAGGTTACGCTGCTAATACCGCATATCTTAAAGAGAAAGCAAGGAGATCGAA 479  
 444 AAAAGGTTACGCTGCTAATACCGCATATCTTAAAGAGAAAGCAAGGAGATCGAA 503  
 480 CTAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539  
 504 CTAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 563  
 540 TAAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599  
 564 TAAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 623  
 600 TAAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659  
 624 CTAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 683  
 660 GTGAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719  
 684 GTGAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 743  
 720 GTGAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779  
 744 GTGAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803  
 780 CTAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839  
 804 CTAAGCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 863

840 AAGTTGACCGGCTGGCGGCGGAGTACGCTGCGAAGATTTAAATCTGAAGATTTGACGGGCGC 899  
 864 AAGTTGACCGGCTGGCGGCGGAGTACGCTGCGAAGATTTAAATCTGAAGATTTGACGGGCGC 923  
 900 CGCACAAGCGGCTGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959  
 924 CGCACAAGCGGCTGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 983  
 960 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016  
 984 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1043  
 1017 GCTGAT 1076  
 1044 GCTGAT 1103  
 1077 AACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136  
 1104 AACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1162  
 1137 ACCGAGGAGGAGTGGGAGTGAACGTCACTGATGATGATGATGATGATGATGATGATGATGAT 1196  
 1163 ACCGAGGAGGAGTGGGAGTGAACGTCACTGATGATGATGATGATGATGATGATGATGATGAT 1222  
 1197 GTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256  
 1223 GTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282  
 1257 GCGCTGAT 1316  
 1283 GCGCTGAT 1342  
 1317 TCGCGGAT 1376  
 1343 TCGCGGAT 1402  
 1377 CCATGAGAGTGGGCTTCAACGAAGCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 1436  
 1403 CCATGAGAGTGGGCTTCAACGAAGCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 1462  
 1437 GTGAGATTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1497  
 1463 GTGAGATTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1531

RESULT 4  
 US-09-342-579-1  
 Sequence 1, Application US/09342579  
 Patent No. 6187569  
 GENERAL INFORMATION:  
 APPLICANT: BRAMUCCI, MICHAEL G.  
 APPLICANT: MCCUTCHEN, CAROL M.  
 APPLICANT: NAGARAJAN, VASANTHA  
 APPLICANT: THOMAS, STUART M.  
 TITLE OF INVENTION: Microbial Production of Terephthalic Acid and  
 TITLE OF INVENTION: Isophthalic Acid  
 FILE REFERENCE: CL-1257  
 CURRENT APPLICATION NUMBER: US/09/342,579  
 CURRENT FILING DATE: 1999-06-29  
 EARLIER APPLICATION NUMBER: 60/091,645  
 EARLIER FILING DATE: JULY 2, 1998  
 NUMBER OF SEQ ID NOS: 4  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 1  
 LENGTH: 1451  
 TYPE: DNA  
 ORGANISM: Burkholderia sp.  
 US-09-342-579-1

Query Match 76.0%; Score 1108; DB 3; Length 1451;  
 Best Local Similarity 87.2%; Pred. No. 0;  
 Matches 1252; Conservative 0; Mismatches 175; Indels 9; Gaps 3;

QY 11 GCGCGCATGCTTTACATGACAGTCCGACGCGAGTCTTGATCTGGTGGCA 70  
DB 11 GCGCGCATGCTTTACATGACAGTCCGACGCGAGTCTTGATCTGGTGGCA 70  
QY 71 GTGGCGAGCGGGTGAATTCATCGGAACTGTCAGAAAGAGGGGGTTACGCAATCGA 130  
DB 71 GTGGCGAGCGGGTGAATTCATCGGAACTGTCAGAAAGAGGGGGTTACGCAATCGA 130  
QY 131 AGATGCTAATATCCGATATCTTAAGAGGAAAGCAGGGGATGGAAGACCTTGGCCG 190  
DB 131 AGCGGATTAATATCCGATATCTTAAGAGGAAAGCAGGGGATGGAAGACCTTGGCCG 190  
QY 191 TTTTGGAGCGCGGATGTCTGATTAAGTGTGGGGTAAAGGCTTACCAAGGGGAGC 250  
DB 191 TTTTGGAGCGCGGATGTCTGATTAAGTGTGGGGTAAAGGCTTACCAAGGGGAGC 250  
QY 251 ATCAAGTATGTTGCTGAGAGAGACGACGACCTGGGACTGAGACAGGCCCAATC 310  
DB 251 ATCTGATCTGCTGAGAGAGACGACGACCTGGGACTGAGACAGGCCCAATC 310  
QY 311 CTACGGGAGGCGACGATGGGGAAATTTGACAAATGGGCGCAAGCTGATCAGCAATGCC 370  
DB 311 CTACGGGAGGCGACGATGGGGAAATTTGACAAATGGGCGCAAGCTGATCAGCAATGCC 370  
QY 371 GCGTGAAGTGAAGAGGCTTCCGGTGTAAAGCTCTTTCAGTCCAGAGAAAGTTACG 430  
DB 371 GCGTGTGTAAAGAGGCTTCCGGTGTAAAGCACTTTGTCCGAAAGAAATCTCTG 430  
QY 431 GTAAATATCTGATCATGACGCTATTCAGACAGAAAGACCGGCTTAATCTGTCGAC 490  
DB 431 GTTAATATCTCGGGGAGATGACCGTAAAGGAAATTAAGACCGGCTTAATCTGTCGAC 490  
QY 491 CAGCCCGGTAATACGATGGGTCGACGCTTAATTCGGAATTAATCGGGCGTAAGGGTGC 550  
DB 491 CAGCCCGGTAATACGATGGGTCGACGCTTAATTCGGAATTAATCGGGCGTAAGGGTGC 550  
QY 551 CAGCGCGCTTTGTAACTGATGATGAAATCCCGGGCTTAATCTGGAATTCGCTTGA 610  
DB 551 CAGCGCGCTTTGTAACTGATGATGAAATCCCGGGCTTAATCTGGAATTCGCTTGA 610  
QY 611 ACTACAGGCTAGAGTGGCAGAGGGAGGTGAAATTCATGTTGTAGCAATGCAATCCGT 670  
DB 611 ACTGCAAGGCTAGAGTGGCAGAGGGAGGTGAAATTCATGTTGTAGCAATGCAATCCGT 670  
QY 671 AGAGATATGGAAGAAATCGATGGCGAAGGCGCTCTGGGTTAACTGACGCTCATG 720  
DB 671 AGAGATATGGAAGAAATCGATGGCGAAGGCGCTCTGGGTTAACTGACGCTCATG 720  
QY 731 CACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATTCACGCTTAAACGATG 790  
DB 731 CACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATTCACGCTTAAACGATG 790  
QY 791 TCAACTAGTGTGGGCTTAAATTAAGGCTTGTGTAACGAGCTTAACGCTGAAGTTGACCG 850  
DB 791 TCAACTAGTGTGGGCTTAAATTAAGGCTTGTGTAACGAGCTTAACGCTGAAGTTGACCG 850  
QY 851 CTGGGAGTACGCTGCGAAGATTAATACTCAAGAGATTGACGGGGACCGGCAACAGCG 910  
DB 851 CTGGGAGTACGCTGCGAAGATTAATACTCAAGAGATTGACGGGGACCGGCAACAGCG 910  
QY 911 TGGATTAATGATTAATTCGATGCAACGCAAAACCTTAACCTTACCTTGAATGAC 970  
DB 911 TGGATTAATGATTAATTCGATGCAACGCAAAACCTTAACCTTGAATGAC 970  
QY 971 GAATTTTCTAGAGATTAATTAAGTCT--TCGGAGCGCTTAACAGAGTGTGATGCT 1027  
DB 971 GAATTTTCTAGAGATTAATTAAGTCT--TCGGAGCGCTTAACAGAGTGTGATGCT 1027  
QY 1028 GTGCTGAGCTGCTGCTGAGAGTGTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1087  
DB 1028 GTGCTGAGCTGCTGCTGAGAGTGTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1087  
QY 1030 GTGCTGAGCTGCTGCTGAGAGTGTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1089  
DB 1030 GTGCTGAGCTGCTGCTGAGAGTGTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1089

QY 1088 TTAATTCGATCATTTTGGTGGGCACTTTAATGAGACTGCCGGTGACAAACCGAGAG 1147  
DB 1090 TTAATTCGATCATTTTGGTGGGCACTTTAATGAGACTGCCGGTGACAAACCGAGAG 1144  
QY 1148 GTGGGAGTACGCTCAAGTCTTCATGATGCGCTTAATGAGGAGGCTTGAACGTAATCAATG 1207  
DB 1145 GTGGGAGTACGCTCAAGTCTTCATGATGCGCTTAATGAGGAGGCTTGAACGTAATCAATG 1204  
QY 1208 GCGCGTACAGAGGGTTCGCAACCCCGAGGGGAGACTAATCTCAGAAAGCGGCTGAT 1267  
DB 1205 GTGCGTACAGAGGGTTCGCAACCCCGAGGGGAGACTAATCTCAGAAAGCGGCTGAT 1264  
QY 1268 CCGGATTCGAGTCTGCACTGCACTCCGGAAGTGGAAATCGCTAATTCGCGATCAG 1327  
DB 1265 CCGGATTCGAGTCTGCACTGCACTCCGGAAGTGGAAATCGCTAATTCGCGATCAG 1324  
QY 1328 CATGTCGCGGTGAATACGTTCCGGGCTTGTACACACCGCCGTCACATGGGAGTG 1387  
DB 1325 CATGTCGCGGTGAATACGTTCCGGGCTTGTACACACCGCCGTCACATGGGAGTG 1384  
QY 1388 GGTTCACCAAGAGAGTATGCTTAACGTAAGAGGGCGCTTCCAGGCTAGAT 1443  
DB 1385 GGTTCACCAAGAGAGTATGCTTAACGTAAGAGGGCGCTTCCAGGCTAGAT 1440

RESULT 5  
US-09-617-854A-1  
; Sequence 1, Application US/09617854A  
; Patent No. 6461840  
; GENERAL INFORMATION:  
; APPLICANT: BRAMUCCI, MICHAEL G.  
; APPLICANT: MCCUTCHEN, CAROL M.  
; APPLICANT: NAGARAJAN, VASANTHA  
; APPLICANT: THOMAS, STUART M.  
; TITLE OF INVENTION: Microbial Production of Terephthalic Acid and  
; TITLE OF INVENTION: Isophthalic Acid  
; FILE REFERENCE: CL-1257  
; CURRENT APPLICATION NUMBER: US/09/617,854A  
; CURRENT FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: 60/091,645  
; PRIOR FILING DATE: JULY 2, 1998  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 1  
; LENGTH: 1451  
; TYPE: DNA  
; ORGANISM: Burkholderia sp.  
US-09-617-854A-1

Query Match 76.0%; Score 1108; DB 3; Length 1451;  
Best Local Similarity 87.2%; Pred. No. 0;  
Matches 1252; Conservative 0; Mismatches 175; Indels 9; Gaps 3;  
QY 11 GCGGATGCTTTACATGACAGTCCGACGCGAGTCTTGATCTGGTGGCA 70  
DB 11 GCGGATGCTTTACATGACAGTCCGACGCGAGTCTTGATCTGGTGGCA 70  
QY 71 GTGGCGAGCGGGTGAATTCATCGGAACTGTCAGAAAGAGGGGGTTACGCAATCGA 130  
DB 71 GTGGCGAGCGGGTGAATTCATCGGAACTGTCAGAAAGAGGGGGTTACGCAATCGA 130  
QY 131 AGATGCTAATATCCGATATCTTAAGAGGAAAGCAGGGGATGGAAGACCTTGGCCG 190  
DB 131 AGCGGATTAATATCCGATATCTTAAGAGGAAAGCAGGGGATGGAAGACCTTGGCCG 190  
QY 191 TTTTGGAGCGCGGATGTCTGATTAAGTGTGGGGTAAAGGCTTACCAAGGGGAGC 250  
DB 191 TTTTGGAGCGCGGATGTCTGATTAAGTGTGGGGTAAAGGCTTACCAAGGGGAGC 250  
QY 251 ATCAAGTATGTTGCTGAGAGAGACGACGACCTGGGACTGAGACAGGCCCAATC 310  
DB 251 ATCTGATCTGCTGAGAGAGACGACGACCTGGGACTGAGACAGGCCCAATC 310



QY	181	GACCTTCGCTTTTGGAGCGCGCGCATCTGCAATTAGTACTTGGGGTAAAGGCTAC	240
Db	178	GGCTTCGCGTACTGGAGCGCGCGCATATCAAGATTAGTATGTTGGTGGGTTAAAGGCTTAC	237
QY	241	CAAGCGCACAGTCAGTAGTTGTTCTGAGAGACGACCAAGCCACACTGGGACTGAGACACG	300
Db	238	CAAGCCACGATCTGTAGCTGTCTTGAGAGGACGACAGCTGGGAAATTTTGGACAATGGCCGAAGCTGTATC	357
QY	301	GCCGAGACTCTCTACGGGAGGACGACGTGGGAAATTTTGGACAATGGCCGCAAGCTGTATC	360
Db	298	GCCGAGACTCTCTACGGGAGGACGACGTGGGAAATTTTGGACAATGGCCGAAGCTGTATC	357
QY	361	CAGCAATGCCCGCTGAGTGAAGAAAGGCTTCGGGTTGTAAGCTCTTCACTGAGAGAGA	420
Db	358	CAGCAATGCCCGCTGAGAGAAAGGCTTCGGGTTGTAAGCTCTTGTACAGAACGA	417
QY	421	AAAGGTTACGGTAAATAATCGTACATGACGCTATCGACAGAAAGACACGGGCTTAC	480
Db	418	AAAGGCTCTGTTAATACCTGGGGCTATGACGGTTACTGTAAAGTAAGCACGGGCTTAC	477
QY	481	TACGTGCCAGACCGCGGGTAATACGTAGGGGTGCAAGCCTTAATCGGAATTTACTGGGCGT	540
Db	478	TACGTGCCAGACCGCGGGTAATACGTAGGGGTGCAAGCCTTAATCGGAATTTACTGGGCGT	537
QY	541	AAAGGTGCCGAGCGCGCTTGTGTAGTCAAGTGAATCCCGGGGCTTAACTCGGGAAAT	600
Db	538	AAAGCGTGGCAGCGCGGTTTGTGAAGCCGATGTGAATCCCGGGGCTCAACTGGGAAC	597
QY	601	TGCGTTTGAATCTACAAGCTAGAGTGTGCGACAGGAGGTGGAATTCATGTGTACAG	660
Db	598	TGCATTTGTGACTGCAAGGCTGAGATGGCGGACAGGGGGAAATGGAATTCGGGTGTACAG	657
QY	661	TGAATGCGTAGAGATATGAAGAACATCGATGCGGAAGGACGCTCTCGGTTTACAAT	720
Db	658	TGAATGCGTAGATATGCGGAGAGAACACCGATGCGGAAGGCAATCCCTCGGGCTGCACT	717
QY	721	GACGCTCATGCACAAAGCGTGGGAGAGAAACAGAAATTAGTATACCCTGTATGTCAGAGCC	780
Db	718	GACGCTCATGCACGAAGCGTGGGAGAGAAACAGAAATTAGTATACCCTGTATGTCAGAGCC	777
QY	781	CTAAACGATGCAACTAGTTGTGGGCTTATTAAGGCTTGTGTTAAGCAAGCTAACGCGTGA	840
Db	778	CTAAACGATGCAACTAGTTGTGGGCTTATTAAGGCTTGTGTTAAGCAAGCTAACGCGTGA	837
QY	841	AGTTGACCGGCTGGGAGTACGCTGCAAGATTAACCTCAAGGAATTTGACGGGACCC	900
Db	838	AGTTGACCGGCTGGGAGTACGCGCGCAAGGTTGAATCTCAAGGAATTTGACGGGACCC	897
QY	901	GCACAAGCGGTGTGATTATGTGATTAAATTCGATGCAACGCGAAAAAACCCTTACCTACCTT	960
Db	898	GCACAAGCGGTGTGATGTGTGTTAAATTCGATGCAACGCGAAAAAACCCTTACCTACCTT	957
QY	961	GACATGTAGCGAAATTTTCTAGAGATAGATTAGTGC---TTGGGGAACGCTATACAGAGTG	1017
Db	958	GACATGTAGCGAAATTTCCGAAAGATTTTGGAAAGTGTCTGTAAGAGAACCTGACACAGGTG	1017
QY	1018	CTGCAATGGCTGTCTCAAGCTCGTGTCTGTAGATGTGTGGGTTAAAGTCCGCAACGAGCGCA	1077
Db	1018	CTGCAATGGCTGTCTCAAGCTCGTGTCTGTAGATGTGTGGGTTAAAGTCCGCAACGAGCGCA	1077
QY	1078	ACCCTTGTCAATTAATTTGCAATCATTTTGGTTGGGCACTTTAATGACATGCGCGGTGACAA	1137
Db	1078	ACCCTTGTCAATTAATTTGCAATCATTTTGGTTGGGCACTTTAATGAGACTGCGGTGACAA	1132
QY	1138	CCGAGGAAGAGTGGGAGTAGAGCTCAAGTGTCTCATAGGCGCTTATGGGTAGGGCTTACACG	1197
Db	1133	CCGAGGAAGAGTGGGAGTAGAGCTCAAGTGTCTCATAGGCGCTTATAGGTGGGCTACACAG	1192
QY	1198	TAAATCAATGGCGGTACAGAGGTTGCCAACCCGCGAGGGGAGCTATCTCAGAAAGC	1257
Db	1193	TCAATCAATGGCGGTACAAAGGTAGCCAAACCCGCGAGGGGAGCAATCCCATTAAGAC	1252
QY	1258	GGGTGTAGTCCGGAATCGGAGTCTGCACTCGATCTCGTGAAGTGGGAATCGTATAGTAT	1317

[illegible]



```
Db 1033 AAGGCTCTCTTAAATACAGGGGGCATATGACGATACCTTAAGAAATTAACCGGCTAAC 974
Qy 481 TACGTCCAGCAGCCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAAATTACTGGCGGT 540
Db 973 TACGTCCAGCAGCCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAAATTACTGGCGGT 914
Qy 541 AAGGGTCCGAGGGCGCTTTCTAAGTACGATGTGAATCCCGGGGCTTAACTGGGAAT 600
Db 913 AAGGGTCCGAGGGCGCTTTCTAAGTACGATGTGAATCCCGGGGCTTAACTGGGAAT 854
Qy 601 TCGCTTTGAAACTACAGAGCTAGTGTGCGAGAGGAGTGAATTCATGTAGCAG 660
Db 853 TCGCTTTGAGTGTGAGAGCTAGTGTGCGAGAGGAGTGAATTCATGTAGCAG 794
Qy 661 TGAATGCGTAGAGATATGGAAGAACTATGATGCGAGAGCGCTCTGGGTTAACT 720
Db 793 TGAATGCGTAGATATGCGAGAGAACTATGATGCGAGAGCGCTCTGGGCTGCACT 734
Qy 721 GACGCTATGCAAGAAAGCTGGGAGCAAGATTAATCCCTGTGATCCAGCGC 780
Db 723 GACGCTATGCAAGAAAGCTGGGAGCAAGATTAATCCCTGTGATCCAGCGC 674
Qy 781 CTAACGATGTCACTAGTTGTTGGGCTTATTAAGCTTGTGACGAAGCTAACCGTGA 840
Db 673 CTAACGATGTCACTAGTTGTTGGGAATTCATTTTCTGATGAAGAACTAACCGTGA 614
Qy 841 AGTTGACCGCTGGGAGATACGCTGCAAGATTAATACTCAAGAAATTAAGCGGAGCC 900
Db 613 AGTTGACCGCTGGGAGATACGCGCGCAAGGTTAAACTCAAGAAATTAAGCGGAGCC 554
Qy 901 GCACAAAGGGTGGATTAATGTGATTAATTCGATGCAAGCGGAAATCCCTTACCTCCT 960
Db 553 GCACAAAGGGTGGATGTGTTAATTCGATGCAAGCGGAAATCCCTTACCTCCTT 494
Qy 961 GACATGTAGCGAATTTCTAGATAGATAGTAACTGCT---TCGGAAACGCTAACAGAGTG 1017
Db 493 GACATGTAGCGAATCTCTGAGAGAGCGAGAGTGTCCGAAAGAGCGCTAACAGAGTG 434
Qy 1018 CTGCATGCTGTGCTGACGCTGTGTGATGATGTTGGTTAAGTCCGCAACGAGCGCA 1077
Db 433 CTGCATGCTGTGCTGACGCTGTGTGATGATGTTGGTTAAGTCCGCAACGAGCGCA 374
Qy 1078 ACCCTTGCATTAATTTGCAATCATTTGTTGGGCACTTAAAGACATGCGCGTGAACA 1137
Db 373 ACCCTTGCATTAATTTGCTACGAAA-----GGGCACTCTTAATGGAATGCGGTAACA 319
Qy 1138 CCGGAGAAAGTGGGAGATGACGTCAGATCCTCATGGCCCTTAATGGGTAGAGGCTTCA 1197
Db 318 CCGGAGAAAGTGGGAGATGACGTCAGATCCTCATGGCCCTTAATGGGTAGAGGCTTCA 259
Qy 1198 TAATACAAATGGCGCTACAGAGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAAG 1257
Db 258 TCATACAAATGGCTGTGTAACAAGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGAAAG 199
Qy 1258 GCGTGTAGTCCGATTCGGAATGCGAATCTCAATCCGGAAGTGGGAATGCGTGTAT 1317
Db 198 GCGTGTAGTCCGATTCGGAATGCGAATCTCAATCCGGAAGTGGGAATGCGTGTAT 139
Qy 1318 GCGGATCAGCATGTCGCGGTGAATACGTTCCGGGTCTTGTACACACGCGCCGTCAAC 1377
Db 138 GCGGATCAGCATGTCGCGGTGAATACGTTCCGGGTCTTGTACACACGCGCCGTCAAC 79
Qy 1378 CATGGAGTGGGTTTACACAGAGAGGATGTCTAAACGTAAGAGAGGCGCTTGCACGG 1437
Db 78 CATGGAGAGGGGTTTGCACAGAGAGTATGCTTAACCGCAAGAGAGGCGATTAACA 19
Qy 1438 TGAATTCATGACTGGG 1455
Db 18 CAGGTTGCTGACTGGG 1
```

```
US-08-642-229A-3
; Sequence 3, Application US/08642229A
; Patent No. 5874291
; GENERAL INFORMATION:
; APPLICANT: Herwig, Russell P.
; APPLICANT: Beilefeldt, Angela R.
; APPLICANT: Stensel, H. David
; APPLICANT: Strand, Stuart E.
; TITLE OF INVENTION: Degradation of Environmental Toxins by a
; FILAMENTOUS BACTERIUM
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson and Kindness PLLC
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: WA 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,229A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/246,865
; FILING DATE: 20-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheinberg, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: UOFW19233
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682 8100
; TELEFAX: (206) 224 0779
; TELEX: 4938023
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1455 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
; DESCRIPTION: "16S ribosomal DNA"
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Unknown. Possibly new species
; STRAIN: A-1
; US-08-642-229A-3
Query Match 75.5%; Score 1100.2; DB 2; Length 1455;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1269; Conservative 1; Mismatches 180; Indels 12; Gaps 5;
Qy 1 ATTGAACGCTGCGCGCATGCTTTACACATGCAAGTGCAGGAGGAGGAGGATGCTGCAT 60
Db 1 ATTGAACGCTGCGCGCATGCTTTACACATGCAAGTGCAGGAGGAGGAGGATGCTGCAT 58
Qy 61 CTGGTGGGAGTGGGCGAGCGGGTGTGTAATGATCGGAACGTATCCAGAAAGGAGGAGTA 120
Db 59 CTGGTGGGAGTGGGCGAGCGGGTGTGTAATGATCGGAACGTATCCAGAAAGGAGGAGTA 118
Qy 121 ACGATCGAAAGATGTCTAATACCGCATATATCTTAAGAGAGGAGGAGGATCGAAA 180
Db 119 ACTACTCGAAAGAGTACTAATACCGCATATATCTTAAGAGAGGAGGAGGATCGAAA 178
Qy 181 GACCTTGCTTTTGGAGCGGCGCGATGTCTGATTAAGTGTGGTGGGCTTAAAGGCTTAC 240
Db 179 GACCTTGCTTACGACGAGCGGCGCTGTGCGAATTAAGTGTGGTGGGCTTAAAGCTTAC 238
```







Db 1120 AAGGCTCTTCTTAATAAAGGCAATGACGGTACCGTAAGATTAAACACCGGCTAAC 1061  
Qy 481 TACGTCAGCAGACCGCGGTATATAGTAGGGGTGCAAGCGTTAATCGAATTACTGGGCGT 540  
Db 1060 TACGTCAGCAGACCGCGGTATATAGTAGGGGTGCAAGCGTTAATCGAATTACTGGGCGT 1001  
Qy 541 AAAAGGTCGCGACGCGGCTTTGTAAATCATAGTGAATCCCGGGCTTAACCTGGGAAT 600  
Db 1000 AAAAGGTCGCGACGCGGCTTTGTAAAGACAGTGTGAATCCCCGGGCTTCACCTGGGAAC 941  
Qy 601 TGGCTTTGAAAATTACAAAGCTGAGTGGCGAGAGGGAGTGAATTCACGTGTAGCAG 660  
Db 940 TGGCAATTGAGCTGCAAGGCTGAGTACGCGACAGAGGGAGTGAATTCGCGGTAGCAG 881  
Qy 661 TGAATGCGTAGAGATATAGAAACATGATGCGAAGGCGACGCTCTCGGCTTAACACT 720  
Db 880 TGAATGCGTAGATATATGCGAGGAACACCGATGCGAAGGCGAGTCCCTCGGCGCTGTACT 821  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTGATACCTCGTGTAGTCCAGCGC 780  
Db 820 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTGATACCTCGTGTAGTCCAGCGC 761  
Qy 781 CTAAACGATGTCACTAGTTGTGGGCTTATTAAGCTGTGAAGCAAGCTAACGCGTGA 840  
Db 760 CTAAACGATGTCACTAGTTGTGGGATTTATTTCTTGAATGACAACTAACGCGTGA 701  
Qy 841 AGTTGACGCGCTGGGGAGTACGGTGCAGAGATTAAACTCAAGAAATGACGGGGACCC 900  
Db 700 AGTTGACGCGCTGGGGAGTACGGCGCGCAAGTTAAACTCAAGAAATGACGGGGACCC 641  
Qy 901 GCAACAAGCGGTGATTAATGATTAATTCATGCAACCGCAAAACCTTAACCTTACCTT 960  
Db 640 GCAACAAGCGGTGATTAATGATTAATTCATGCAACCGCAAAACCTTAACCTTACCTT 581  
Qy 961 GACATGTACGCAATTTTCTAGATAGATTAGTGTCT--TCGGGAACGCTAACAGAGT 1017  
Db 580 GACATGTATGAGATCCCGAGAGATGTGGAGTGTGCGAAGAGACCAATACACAGGAG 521  
Qy 1018 CTGCAATGAGCTGCTGAGCTGCTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1077  
Db 520 CTGCAATGAGCTGCTGAGCTGCTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 461  
Qy 1078 ACCCTTGTCAATTAATTCATCAATCTTTGTTGGGCACTTTAATGACACTGCGGTGACAA 1137  
Db 460 ACCCTTGTCAATTAATTCATCAATCTTTGTTGGGCACTTTAATGACACTGCGGTGACAA 406  
Qy 1138 CCGGAGGAAGGTGGGATGACGTCAGTCTCTCATGTCCTTAATGGGTAGGGCTTCAACAG 1197  
Db 405 CCGGAGGAAGGTGGGATGACGTCAGTCTCTCATGTCCTTAATGGGTAGGGCTTCAACAG 346  
Qy 1198 TAATACATAGGCGCGTACAGAGGTTGCCAACCCGAGGGGGAGCTAATCTCAGAAAGC 1257  
Db 345 TAATACATAGGCGCGTACAGAGGCGAGCAAGCGCGAGGTGAAGCAATCCATTAAG 286  
Qy 1258 GCGTCGTAGTCCGATCGAGTCTGCACTGACTCGTGAAGTCCGAATCGCTAGTAAT 1317  
Db 285 GCGTCGTAGTCCGATCGAGTCTGCACTGACTCGTGAAGTCCGAATCGCTAGTAAT 226  
Qy 1318 CCGGATCAGCATGTGCGCGGTGAATAGTTCCTCCCGGCTTTGTACACACCGCCCGTCAAC 1377  
Db 225 CCGGATCAGCATGTGCGCGGTGAATAGTTCCTCCCGGCTTTGTACACACCGCCCGTCAAC 166  
Qy 1378 CATGGGAGTGGGTTTTCACAGAAAGAGGTAGTCTAACCCGTAAGAGAGGCGCTTGCCACG 1437  
Db 165 CATGGGAGTGGGTTTTCACAGAAAGAGGTAGTCTAACCCGTAAGAGAGGCGCTTGCCACG 106  
Qy 1438 TGAATTCATGACTGGGGTG 1457  
Db 105 TGGGTTTCTGATCTGGGGTG 86

RESULT 11  
US-09-735-567-7

Sequence 7, Application US/09735567  
Patent No. 6608190  
GENERAL INFORMATION:  
APPLICANT: Bramucci, Michael  
TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
FILE OF INVENTION: Industrial Wastewater Bioreactors  
FILE REFERENCE: BC1033 US NA  
CURRENT FILING DATE: US/09/735.567  
PRIOR FILING DATE: 2000-12-13  
PRIOR APPLICATION NUMBER: 60/171,140  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 7  
LENGTH: 1454  
TYPE: DNA  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Brachyomonas  
OTHER INFORMATION: Nucleotide sequence of the 16S rDNA corresponding to strain RA9  
US-09-735-567-7  
Query Match 74.1%; Score 1080.2; DB 3; Length 1454;  
Best Local Similarity 85.6%; Pred. No. 0;  
Matches 1251; Conservative 0; Mismatches 198; Indels 12; Gaps 4;  
Qy 2 TTGAACGCTGGCGGATCTTTACATGCAAGTGAACCGGACGACCGGATCTTGCAATC 61  
Db 1 TTGAACGCTGGCGGATCTTTACATGCAAGTGAACCGGACGACCGGATCTTGCAATC 58  
Qy 62 TGGTGGCGAGTGGCGGAGGGGTGATGATGATGCAAGTATCCGCAAGAGAGGGGTAA 121  
Db 59 TGGTGGCGAGTGGCGGAGGGGTGATGATGATGCAAGTATCCGCAAGAGAGGGGTAA 118  
Qy 122 CGCATGCAAGATGTCTAATACCGCATATCTCTAAGAGAGAAAGAGAGGAG--TCGA 179  
Db 119 CTGGCGCAAGAGCGGATTAATACCGCATGATCCGTGATGAAGAGAGAGGAGACTCGCA 178  
Qy 180 AGACCTTGGCTTTTGAAGCGGCGCATGCTGATTAAGTATGTTGGTGGGTTAAAGCCCTA 239  
Db 179 GGGCTTGGCTTGAAGCGGCGCATGCTGATTAAGTATGTTGGTGGGTTAAAGCCCTA 238  
Qy 240 CCAAGCGCATGCTGATGTTGTTGAGAGAGCAACGACCTGGGATCTGAGACAC 299  
Db 239 CCAAGCGCATGCTGATGTTGTTGAGAGATTAATCGGATCTGAGACAC 298  
Qy 300 GGCAGGACTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 359  
Db 299 GGCAGGACTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 358  
Qy 360 CCAGCAATGCGCGTGAAGTGAAGAGGCTTGGGTTGAAGCTCTTCACTCGAAG 419  
Db 359 CCAGCAATGCGCGTGAAGTGAAGAGGCTTGGGTTGAAGCTCTTCACTCGAAG 418  
Qy 420 AAAAGTTACGTTAATTAATCTGACTGATGACGAGTATGACAGAGAGAGAGAGAGAG 479  
Db 419 AAAAGTTACGTTAATTAATCTGACTGATGACGAGTATGACAGAGAGAGAGAGAGAG 478  
Qy 480 CTACGTCAGCAGCGCGGTATATAGTGAAGGCTTAATCGAATTACTGGGCG 539  
Db 479 CTACGTCAGCAGCGCGGTATATAGTGAAGGCTTAATCGAATTACTGGGCG 538  
Qy 540 TAAAGGTCGAGGCGGCTTTGTAAAGTGAATGCAATGCAATGCAATGCAATGCAATGCA 599  
Db 539 TAAAGGTCGAGGCGGCTTTGTAAAGTGAATGCAATGCAATGCAATGCAATGCAATGCA 598  
Qy 600 TTGCTTTAAACTACAGGCTGATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659  
Db 599 TTGCTTTAAACTACAGGCTGATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658  
Qy 660 GTGAATGCGTGAAGTATGAGAAACATGATGCGAAGGAGGAGGAGGAGGAGGAGGAGGAG 719  
Db 659 GTGAATGCGTGAAGTATGAGAAACATGATGCGAAGGAGGAGGAGGAGGAGGAGGAGGAG 718

```
OY 720 TGACCCCTCATGACGAAAGCGTGGGAGCAACAGAGTTAGTACCCTTGATGTCACAGC 779
Db 719 TGACCCCTCATGACGAAAGCGTGGGAGCAACAGAGTTAGTACCCTTGATGTCACAGC 778
OY 780 CCTAAACGATGCACTAGTTGTTGGGCTTATAGGCTTGATTAAGCAAGCAAGCAAGCGCTG 839
Db 779 CCTAAACGATGCACTAGTTGTTGGGCTTATAGGCTTATAGGCAAGCAAGCAAGCGCTG 838
OY 840 AAGTTACCGCTGGGAGATGCGGTGCAAGATTAAACTCAAGAAATTGACGGGAGC 899
Db 839 AAGTTACCGCTGGGAGATGCGGTGCAAGATTAAACTCAAGAAATTGACGGGAGC 898
OY 900 CGCACAAAGCGGTGATTTATGTGATTTAATTGATGCAAGCGAAAAACCTTACCTACCT 959
Db 899 CGCACAAAGCGGTGATTTATGTGATTTAATTGATGCAAGCGAAAAACCTTACCTACCT 958
OY 960 TGACATGTAGAGCAATTTCTAGAGATGATGCTCTGCGGAAAGCTTAACAGAGT 1016
Db 959 TGACATGTAGAGCAATTTCTAGAGATGATGCTCTGCGGAAAGCTTAACAGAGT 1018
OY 1017 GCTGATGAGCTGTCGTACAGCTGTCGTGAGATGTTGGGTTAAGTCCCGCAAGAGCGC 1076
Db 1019 GCTGATGAGCTGTCGTACAGCTGTCGTGAGATGTTGGGTTAAGTCCCGCAAGAGCGC 1078
OY 1077 AACCTTGTCAATTAATGTCATCAATTTGGTGGGCACTTTAATGAGACTGCCGGTGACA 1136
Db 1079 AACCTTGTCAATTAATGTCATCAATTTGGTGGGCACTTTAATGAGACTGCCGGTGACA 1133
OY 1137 ACCGAGGAGAGGTGGGAGATGACGTCAAGTCTCATGAGCCCTATGAGGTAGGCTTACAGC 1196
Db 1134 ACCGAGGAGAGGTGGGAGATGACGTCAAGTCTCATGAGCCCTATGAGGTAGGCTTACAGC 1193
OY 1197 GTAAATCAATGCGCGTACAGAGGCTTGGCAACCCGCGAGGGGAGCTAATCTCAAGAA 1256
Db 1194 GTAAATCAATGCGCGTACAGAGGCTTGGCAACCCGCGAGGGGAGCTAATCTCAAGAA 1253
OY 1257 GCGCTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 1316
Db 1254 GCGCTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 1313
OY 1317 TCGCGGATCAGATGTCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCCTACA 1376
Db 1314 TCGCGGATCAGATGTCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCCTACA 1373
OY 1377 CCATGAGAGTGGGTTTACACAGAGCAGGTAGTCTTAACCTTAAGAGAGGCGCTTCCAC 1436
Db 1374 CCATGAGAGTGGGTTTACACAGAGCAGGTAGTCTTAACCTTAAGAGAGGCGCTTCCAC 1433
OY 1437 GTGAGATTCATGACTGGGGTG 1457
Db 1434 GTGAGATTCATGACTGGGGTG 1454
```

```
RESULT 12
US-09-735-567-2/c
; Sequence 2, Application US/09735567
; Patent No. 6608190
; GENERAL INFORMATION:
; APPLICANT: Brannucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE REFERENCE: BCI033 US NA
; CURRENT APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Unknown Organism
```

```
FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Brachyomae
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain NBVE13
US-09-735-567-2
```

```
Query Match 74.1%; Score 1079.6; DB 3; Length 1467;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 1251; Conservative 0; Mismatches 199; Indels 12; Gaps 4;
```

```
OY 1 ATTGAAGCTGCGCGCATGCTTTACATGCAATGCAATCGGACGACAGATGCTTGCAT 60
Db 1465 ATTGAAGCTGCGCGCATGCTTTACATGCAATGCAATCGGACGACGCTGCTTGCAT 1408
OY 61 CTGGTGGGAGTGGCGGAGCGGATGATATGCAATCGGAAAGCTATCCAGAAAGGAGGATA 120
Db 1407 ATGGCGGAGATGGCGAAAGCGGATGATATGCAATCGGAAAGCTATCCAGAAAGGAGGATA 1348
OY 121 ACGCATGAAAGATGTCATATACCGCATATATCTTAAGAGAGAAAGCAGAGGA -TGA 178
Db 1347 GCTGGCGAAAGCGGATTAATACCGCATATATCTTGGATGAAAGCAGAGGAGCTGCA 1288
OY 179 AAGACTTGGCGCTTTGGAGCGCCGATGTCATATGCTAATGTTGTTGGGTTAAAGCCT 238
Db 1287 AGGACTTGGCGCTTGGAGCGCCGATGTCATATGCTAATGTTGTTGGGTTAAAGCCT 1228
OY 239 ACCAAGCGAAGATCAGATGATGTTGTTGAGAGAGAGCAAGCGACACCTGGGACTGAGCA 298
Db 1227 ACCAAGCGGATGATGATGTTGTTGAGAGAGAGCAAGCGGATGTTGAGCAATGAGGAGCTGA 1168
OY 299 CGGCGCAGACTCTTACGAGAGAGCAGAGTGGGAAATTTGGACAAATGGCGCAAGCTGA 358
Db 1167 CGGCGCAGACTCTTACGAGAGAGCAGAGTGGGAAATTTGGACAAATGGCGCAAGCTGA 1108
OY 359 TCCAGCAATGCGCGTGAATGAGAGAGCCTTGGGTTGTTAAAGCTTTTCACTGAGAA 418
Db 1107 TCCAGCAATGCGCGTGAATGAGAGAGCCTTGGGTTGTTAAAGCTTTTCACTGAGAA 4048
OY 419 GAAAGGTTTACGTTAAATTAATGTCATCAATGAGGATGCAAGAGAAAGCAGGCTA 478
Db 1047 GAAAGGCTCTTTTAAATTAAGAGAGGCAATGAGGATGCAAGAGAAATTAAGAGGCTA 988
OY 1047 GAAAGGCTCTTTTAAATTAAGAGAGGCAATGAGGATGCAAGAGAAATTAAGAGGCTA 988
OY 479 ACTAGTGCAGAGCGCGGTAATAGTAGGAGTGAAGCGTTAATCGAAATTAAGTGGG 538
Db 987 ACTAGTGCAGAGCGCGGTAATAGTAGGAGTGAAGCGTTAATCGAAATTAAGTGGG 928
OY 539 GTAAAGGTTGCGAGGCGCTTTGTAAGTCAATGTAATCCCGGGCTTAACCTGGGA 598
Db 927 GTAAAGGTTGCGAGGCGCTTTGTAAGTCAATGTAATCCCGGGCTTAACCTGGGA 868
OY 599 ATTGCGTTTGAATCAACAGGCTAGAGTGGCAGAGGAGTGAATTTCCATGTTGAGC 658
Db 867 ACTGCGTTTGAATCAACAGGCTAGAGTGGCAGAGGAGTGAATTTCCATGTTGAGC 808
OY 867 ACTGCGTTTGAATCAACAGGCTAGAGTGGCAGAGGAGTGAATTTCCATGTTGAGC 808
Db 659 AGTGAATGCGTAGAGATATGGAAGAAATGATAGTGGAGGAGCGCTCTGGGTTAACA 718
OY 807 AGTGAATGCGTAGAGATATGGAAGAAATGATAGTGGAGGAGCGCTCTGGGTTAACA 748
Db 719 CTGAGCTCATGCAAGAAAGCTGGGAGCAAAAGATTAATCCCTGGTATTCACG 778
OY 747 CTGAGCTCATGCAAGAAAGCTGGGAGCAAAAGATTAATCCCTGGTATTCACG 688
Db 779 CCTTAACGATGCAATGATGTTGGGCTTATTAAGGCTTGAAGCAAGCTAAGCGGT 838
OY 687 CCTTAACGATGCAATGATGTTGGGCTTATTAAGGCTTGAAGCAAGCTAAGCGGT 628
Db 839 GAAATTGACCGCTGGGAGTACCGGTGCAAGATTAAACTCAAGAAATTTGACGGGAGC 898
OY 627 GAAATTGACCGCTGGGAGTACCGGTGCAAGATTAAACTCAAGAAATTTGACGGGAGC 568
Db 899 CGGACAAAGCGGTGATATGATGATTAATTCATGCAAGCGGAAAAACCTTACCTACCC 958
OY 567 CGGACAAAGCGGTGATATGATGATTAATTCATGCAAGCGGAAAAACCTTACCTACCC 508
```

QY 959 TTGACATGTAGCGAATTTTCTAGATAGATTAGTCT---TCGGAAAGCTTAACACAGG 1015  
DB 507 TTGACATGTATGAAATCCCGCAGAGATGTGGAGTGTCTGCAGAGAGACATTAACACAGG 448  
QY 1016 TGCCTCATGCGCTGTGTGACGCTCGTGTGTGATGTTAGGTTAAGTCCCGCAACGAGG 1075  
DB 447 TGTCTCATGCTGTGTGTGACGCTCGTGTGTGATGTTAGGTTAAGTCCCGCAACGAGG 388  
QY 1076 CAACCTTGTCTAATTAATTCATCATTTGTTGGGCACTTTAATGAGACTGCGGTGACA 1135  
DB 387 CAACCTTGTCTAATTAATTCATCATTTGTTGGGCACTTTAATGAGACTGCGGTGACA 333  
QY 1136 AACCGGAGAGAGTGGAGTGAAGTCAATGCTCATGCTGCTTATGAGGAGGCTTTGACA 1195  
DB 332 AACCGGAGAGAGTGGAGTGAAGTCAATGCTCATGCTGCTTATGAGGAGGCTTTGACA 273  
QY 1196 CGTATATCAATGCGCGGTACAGAGGTTGCCAACCCGCGAGGGGGGCTAATCTGAGAA 1255  
DB 272 CGTATATCAATGCGCGGTACAGAGGTTGCCAACCCGCGAGGGGGGCTAATCTGAGAA 213  
QY 1256 GCGCGTGTAGTCCGGATCGAGTCTGAACCTGCACTCCGTGAAGTCGGAATCGCTAGTA 1315  
DB 212 GCGCGTGTAGTCCGGATCGAGTCTGAACCTGCACTCCGTGAAGTCGGAATCGCTAGTA 153  
QY 1316 ATCGCGGATCAGCATGTGCGCGTGAATACGTTCCGCGTCTTGTACACACCGCGGTGAC 1375  
DB 152 ATCGCGGATCAGCATGTGCGCGTGAATACGTTCCGCGTCTTGTACACACCGCGGTGAC 93  
QY 1376 ACCATGGAGTGGGTTTTCACAGAGAGGTAAGTCTTAACCTTAAGAGAGGCGCTTGCCAC 1435  
DB 92 ACCATGGAGTGGGTTTTCACAGAGAGGTAAGTCTTAACCTTAAGAGAGGCGCTTGCCAC 33  
QY 1436 GGTGAGATTTCATGACTGGGGTG 1457  
DB 32 GGTGAGATTTCATGACTGGGGTG 11

RESULT 13  
US-09-063-898-1  
Sequence 1, Application US/09063898  
Patent No. 6319497  
GENERAL INFORMATION:  
APPLICANT: Casida, Jr., Lester E.  
APPLICANT: Falkinham, III, Joseph Oliver  
APPLICANT: Cain, Cody C.  
TITLE OF INVENTION: NON-OBLIGATE PREDATORY BACTERIUM  
TITLE OF INVENTION: BURKHOLDERIA CASIDAE AND USES THEREOF  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1667 K Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/063,898  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/044,532  
FILING DATE: 23-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mirock, Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 8743-006-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 7909090

TELEFAX: (212) 8699741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Genomic DNA  
US-09-063-898-1  
Query Match 72.8%; Score 1060.8; DB 3; Length 1495;  
Best Local Similarity 86.6%; Pred. No. 0;  
Matches 1264; Conservative 0; Mismatches 182; Indels 14; Gaps 8;  
QY 1 ATTGAACGCTGGCGGCATGCTTTTACA-CATGCAAGTCGAACGGCAGCAGCATGCTTCCA 59  
DB 3 ATATTACGCTGGTTCATGCTCTTACAGCATGCAAGTCGAACGGCAGCAGCATGCTTCCA 62  
QY 60 TCTGTGCGCAGTGGCGGACGGGTGATGATGATCGGAC-GTATCCAGAAAGAGGGGG 118  
DB 63 CTTGTGCGCAGTGGCGGACGGGTGATGATGATCGGACCAATGCTCTGATGTTGGGGA 122  
QY 119 TAAAGCATGAAAAGTGTCTAATACCGCATTAATCTTAAGAGGAGAAACAGGGGATCGA 178  
DB 123 TAGCCCGGAGAAAGCGGATTAATACCGCATTAATCTTAAGAGGAGAAACAGGGGATCGA 182  
QY 179 AAGACCTTGCGCTTTTGAAGCGCGCATGCTGATTAAGTGTGTTGGGTTAAAGGCTT 238  
DB 183 GGGGCTGCGCTTATAGGGTTAGGCTGATGATGATTAAGTGTGTTGGGTTAAAGGCTT 242  
QY 239 ACCAAGGCGACATCAATGTTGTTGAGAGAGACAGACAGCAGCAGCTGAGGAGCTGAGACA 298  
DB 243 ACCAAGGCGACATCAATGTTGTTGAGAGAGACAGACAGCAGCAGCTGAGGAGCTGAGACA 301  
QY 299 GGGGCCAGCTCCTACGGGAGGACAGATGGGGAATTTTGAACAATGGGCGCAAGCTTGA 358  
DB 302 GGGGCCAGCTCCTACGGGAGGACAGATGGGGAATTTTGAACAATGGGCGCAAGCTTGA 361  
QY 359 TCCAGCAATGCCGCGTGAAGTGAAGAGGCTTGGGTTTGAACAATGGGCGCAAGCTTGA 418  
DB 362 TCCAGCAATGCCGCGTGAAGTGAAGAGGCTTGGGTTTGAACAATGGGCGCAAGCTTGA 421  
QY 419 GAAAAGTTACGTTAATATGATGATCAATCAATGATGATGATGATGATGATGATGATGAT 478  
DB 422 GAAATCCTTGTCTAATATATGATGATCAATCAATGATGATGATGATGATGATGATGAT 481  
QY 479 ACTACGTGACACAGCGCGGTAATACGTAAGGAGTGAAGCTTAAATCGGAATTAATCGGAC 538  
DB 482 ACTACGTGACACAGCGCGGTAATACGTAAGGAGTGAAGCTTAAATCGGAATTAATCGGAC 541  
QY 539 GTAAAGGTTGCGCAGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 598  
DB 542 GTAAAGGTTGCGCAGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 601  
QY 599 ATTGCGTTTGAACCTAACAAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 658  
DB 602 ACTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661  
QY 659 AGTGAATGCGTGAAGTATGAGAAACATGATGCGGAAAGCAGCTCTGGGTTAACA 718  
DB 662 AGTGAATGCGTGAAGTATGAGAAACATGATGCGGAAAGCAGCTCTGGGTTAACA 721  
QY 719 CTGAGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGTACCTGTGATGATGATGATGAT 778  
DB 722 CTGAGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGTACCTGTGATGATGATGATGAT 781  
QY 779 CCTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838  
DB 782 CCTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 841  
QY 839 GAAGTTGACCGCTGGGAGTACCGTGGCAAGTTAAATCTCAAGAAATTTGACGGGAGC 898

Db 842 GAAGTTGACCGCGCTGGGAGTAGCGTCCGAAGATTAAAACTCAAGGAATTGACGGGAC 901  
Qy 899 CCGCAAGACGGGTGATATGTGATTAATTCATGCAACCGGAAAACTTACTTACC 958  
Db 902 CCGCAAGACGGGTGATATGTGATTAATTCATGCAACCGGAAAACTTACTTACC 961  
Qy 959 TTGACATGATGGAATTTTCTAGAGATGATTAGGCT---TCGGGAACGGTAAACAGAG 1015  
Db 962 TTGACATGATGGAATTTTCTAGAGATGATTAGGCT---TCGGGAACGGTAAACAGAG 1021  
Qy 1016 TGCTGATGAGCTGTGCTGAGCTGTGTGTGAGATGTTGGGTTAAGTCCCGCAACGAG 1075  
Db 1022 TGCTGATGAGCTGTGCTGAGCTGTGTGTGAGATGTTGGGTTAAGTCCCGCAACGAG 1081  
Qy 1076 CAACCTTGTCTAATTAATTCATCATTTTGTGGCACTTTATGAGACTGCGGTGACA 1135  
Db 1082 CAACCTTGTCTAATTAATTCATCATTTTGTGGCACTTTATGAGACTGCGGTGACA 1136  
Qy 1136 AACCGAGAGAGTGGGAGTAGCGTCAAGTCTCTATGCGCCCTATGGGTAGGGCTTCA 1195  
Db 1137 AACCGAGAGAGTGGGAGTAGCGTCAAGTCTCTATGCGCCCTATGGGTAGGGCTTCA 1195  
Qy 1196 CGTAAATACATGCGCGCTACAGAGGTTGCCAACCGGAGGGGAGCTTAATTCAGAAA 1255  
Db 1196 CGTAAATACATGCGCGCTACAGAGGTTGCCAACCGGAGGGGAGCTTAATTCAGAAA 1254  
Qy 1256 GCGGCTGTGATCGCGATCGAGTTCGCACTCGACTCCGTGAAGTCGGAATCGTACTA 1315  
Db 1255 ACCGATCTGATCGCGATTCGCACTCGACTCGAGTGAATGAAATCGTACTA 1314  
Qy 1316 ATCCGGATTCAGCATGTCGCGGTGAATCGTCCGGGCTGTGACACCGCGCGTAC 1375  
Db 1315 ATCCGGATTCAGCATGTCGCGGTGAATCGTCCGGGCTGTGACACCGCGCGTAC 1374  
Qy 1376 ACCATGGAAGTGGTTTCAACGAGAGTAGTCTAACCGTGAAG-AGGCGCTTGCA 1434  
Db 1375 ACCATGGAAGTGGTTTCAACGAGAGTAGTCTAACCGTGAAG-AGGCGCTTGCA 1434  
Qy 1435 CCGTAGATTCATGACTGGG 1454  
Db 1435 CCGTAGATTCATGACTGGG 1454

RESULT 14  
US-09-985-846-1  
; Sequence 1, Application US/09985846  
; Patent No. 6689357  
; GENERAL INFORMATION:  
; APPLICANT: Casida, Jr., Lester E.  
; Falkinham, III, Joseph Oliver  
; Cain, Cody C.  
; TITLE OF INVENTION: NON-OBLIGATE PREDATORY BACTERIUM  
; BURKHOLDERIA CASIDAE AND USES THEREOF  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Penile & Edmonds LLP  
; STREET: 1667 K Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Pasteo Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/985,846  
; FILING DATE: 06-NO. 6689357-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/063,898  
; FILING DATE: 22-APR-1998

APPLICATION NUMBER: 60/044,532  
FILING DATE: 23-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE//DOCKET NUMBER: 8743-006-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)7909090  
TELEFAX: (212)8699741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Genomic DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-985-846-1

Query Match 72.8%; Score 1060.8; DB 3; Length 1495;  
Best Local Similarity 86.6%; Pred. No. 0;  
Matches 1264; Conservative 0; Mismatches 182; Indels 14; Gaps 8;

1 ATTGAACGCTGGCGGCAATGCTTTTACA-CATGCAAGTGAACGGGAGACGGATCTTGCA 59  
3 ATATTACGCTGGTTCATGCTTTACAGCATGCAAGTGAACGGGAGACGGATCTTGCA 62  
60 TCTGTGGCGAGTGGCGGACGGGTGAATATGATCGGAAC-GTATCAAGAGGGGGG 118  
63 CCGTGTGGCGAGTGGCGGACGGGTGAATATGATCGGAAC-GTATCAAGAGGGGGG 122  
119 TAAAGCATCGAAGATGTGCTTAATCCGCAATATCTTAAAGGAGAAAGCGGGGATCGA 178  
123 TAGCCCGCGGAAACCGGATTAATACCGCATATGATGAGTGAAGAGCGGGGACCTT 182  
179 AAGACCTTGCGCTTTTGAAGCGCGCATGTCTGATTAAGTGAAGTGGGGTAAAGGCT 238  
183 CCGGCTTCGCGCTTATAGGCTTGGCCGATGCTGATTAAGTGAAGTGGGGTAAAGGCT 242  
239 ACCAAGCGAGATCAATGATGTTGGTCTGAGAGAGACGACCGCACTGGAGCTGAGACA 298  
243 ACCAAGCGAGATCAATGATGTTGGTCTGAGAGAGACGACCGCACTGGAGCTGAGACA 301  
299 CCGGCGCAGATCTCTACCGGAGGCGAGCTGGGGAATTTTGAACAATGGCGGCAAGCT 358  
302 CCGGCGCAGATCTCTACCGGAGGCGAGCTGGGGAATTTTGAACAATGGCGGCAAGCT 361  
359 TCCAGCAATGCGCGGTGAGTGAAGAGAGGCTTGGGTTGTAAAGCTCTTTCAGTCGAGAA 418  
362 TCCAGCAATGCGCGGTGAGTGAAGAGAGGCTTGGGTTGTAAAGCTCTTTCAGTCGAGAA 421  
419 GAAAGGTTACGTTAAATATCGTGAATGACGATGACGGTATGACAGAAAGAGACCGGCTA 478  
422 GAAATCCTGGTCTTAATATATAGCGGGGATGACGATACCGGAAGATTAAGCACCGGCTA 481  
479 ACTACGTGCGACGACCGCGGTATATAGTAGGGTGCAGAGCTTAATGCGAATTACTGGGC 538  
482 ACTACGTGCGACGACCGCGGTATATAGTAGGGTGCAGAGCTTAATGCGAATTACTGGGC 541  
539 GTAAAGGTTGCGGAGCGGCTTTGTAAGTCAAGATGTAATCCCGGAGCTTAATCTGGGA 598  
542 GTAAAGGTTGCGGAGCGGCTTTGTAAGTCAAGATGTAATCCCGGAGCTTAATCTGGGA 601  
599 ATTGCGTTTGAACCTAACAGGCTTAGAGTGTGGCAGAGAGAGTGAATTCATGTGTAGC 658  
602 ACTGCAATTTGATCTGGAGGCTTAGAGTGTGGCAGAGAGAGTGAATTCATGTGTAGC 661  
659 AGTAAATTCCTTAAGATATGAAAGAACTATGATGCGGAGAGGCGCTCTCTGGGTAAACA 718  
662 AGTAAATTCCTTAAGATATGAAAGAACTATGATGCGGAGAGGCGCTCTCTGGGTAAACA 721  
719 CTGACGCTCATGACGAAAGGTTGGGAGGCAAAACAGGATTAAGATTAAGTCAAGC 778

```

Db      |||||
722      CTGACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTAGTACCTGTGAGTCCACG 781
Qy      |||||
779      CCTTAAAGATGCACTAGTTGTTGGGCTTATAGGCTGGTAAACGAAGTAAAGCGGT 838
Db      |||||
782      CCTTAAAGATGCACTAGTTGTTGGGATTCATTTCTTATGTAAGTAAAGCGGT 841
Qy      |||||
839      GAAATTGACCGCTGGGAGTACGCTCCAGATTAACTCAAGGAATTGACGGGAGC 898
Db      |||||
842      GAAATTGACCGCTGGGAGTACGCTCCAGATTAACTCAAGGAATTGACGGGAGC 901
Qy      |||||
899      CCGCAAGACGGTGGATATGATGATTAATTCATGACCAAGGAAAACTTAACTTACC 958
Db      |||||
902      CCGCAAGACGGTGGATATGATGATTAATTCATGACCAAGGAAAACTTAACTTACC 961
Qy      |||||
959      TTGACATGATGCAATTTTCTAGAGATGATTAGTCT---TCGGGAACGCTTAACAGC 1015
Db      |||||
962      TTGACATGATGCAATTTTCTAGAGATGATTAGTCT---TCGGGAACGCTTAACAGC 1021
Qy      |||||
1016     TGCTGATGCTGTGCTGAGCTGCTGCTGAGATGTTGGGTTAAAGTCCCGCAACAGAGC 1075
Db      |||||
1022     TGCTGATGCTGTGCTGAGCTGCTGCTGAGATGTTGGGTTAAAGTCCCGCAACAGAGC 1081
Qy      |||||
1076     CAACCTTGTCTTAAATTTGCTCATTTGTTGGGCACTTTAATGATGCTCCGGTGA 1135
Db      |||||
1082     CAACCTTGTCTTAAATTTGCTCATTTGTTGGGCACTTTAATGATGCTCCGGTGA 1136
Qy      |||||
1136     AACCGGAGAGTGGGAGTGAAGTCAATTCCTCAAGGCTTAAAGGAGGCTTCA 1195
Db      |||||
1137     AACCGGAGAGTGGGAGTGAAGTCAATTCCTCAAGGCTTAAAGGAGGCTTCA 1195
Qy      |||||
1196     CGTAATACAAATGGCGGCTACAGAGGTTGCAACCCGCGAGGAGAGTAACTCAGAAA 1255
Db      |||||
1196     CGTAATACAAATGGCGGCTACAGAGGTTGCAACCCGCGAGGAGAGTAACTCAGAAA 1255
Qy      |||||
1196     CGTAATACAAATGGCGGCTACAGAGGTTGCAACCCGCGAGGAGAGTAACTCAGAAA 1255
Db      |||||
1256     GCGCGTGTAGTCCGAGTCCGAGTCTGCACTCGACTCGGTGAAGTGGAAATGCTTA 1315
Qy      |||||
1256     GCGCGTGTAGTCCGAGTCCGAGTCTGCACTCGACTCGGTGAAGTGGAAATGCTTA 1315
Db      |||||
1255     ACCGATCGTAGTCCGAGTCCGAGTCTGCACTCGACTCGGTGAAGTGGAAATGCTTA 1314
Qy      |||||
1316     ATCGCGGATCAGATGCTCCGCGGTGAATAGCTTCCCGGGTCTTGTACACCCGCGTCA 1375
Db      |||||
1315     ATCGCGGATCAGATGCTCCGCGGTGAATAGCTTCCCGGGTCTTGTACACCCGCGTCA 1374
Qy      |||||
1376     ACCATGGAGTGGGTTTCAACAGAGCAGTAGTCTTAACCGTAAAG-AGGCGCTTGGCA 1434
Db      |||||
1375     ACCATGGAGTGGGTTTCAACAGAGCAGTAGTCTTAACCGGCAAGAAACGCTCCCA 1434
Qy      |||||
1435     CGGTGAGATTCACTGAGTGGG 1454
Db      |||||
1435     CGGTGAGATTCACTGAGTGGG 1454

```

```

RESULT 15
US-09-228-184-1
; Sequence 1, Application US/09228184
; Patent No. 6322782
; GENERAL INFORMATION:
; APPLICANT: WALKER, Harrell L.
; APPLICANT: HINGINBOTHAM, Lawrence R.
; TITLE OF INVENTION: CONTROL OF CYANOBACTERIA WITH A BACTERIUM
; FILE REFERENCE: 013243-0007
; CURRENT APPLICATION NUMBER: US/09/228,184
; CURRENT FILING DATE: 1999-01-11
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1540
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURES:
; OTHER INFORMATION: Description of Unknown Organism: Bacterium SG-3;
; OTHER INFORMATION: gram negative; rod-shaped; exhibits flagellate
; OTHER INFORMATION: motility; pathogenic to cyanobacteria and algae;

```

OTHER INFORMATION: yellow colonies on BG-11 medium suppl. with tryptic  
US-09-228-184-1

Query Match 72.3%; Score 1053; DB 3; Length 1540;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 1242; Conservative 0; Mismatches 215; Indels 10; Gaps 5;

```

Qy      1  ATTGAAGCTGCGCGCATGCTTTTACATGCAAGTCCAGACGGCAGACGATGCTTGAT 60
Db      24  AGTGAAGCTGCGCGCATGCTTTTACATGCAAGTCCAGACGGCAGACGATGCTTGAT 83
Qy      61  CT-----GGTGGCAGTGGCGGACCGGCTGATGATGCAATCGGAACGTAATCGAAGAGGG 116
Db      84  ACTGTTGGTGGCAGTGGCGGACCGGCTGATGATGCAATCGGAACGTAATCGGAGTGG 143
Qy      117  GGTAAAGCATGGAAGATGCTTAATACCGATTAATCTTAAGAGAGAAACAGGAGATC 176
Db      144  GATTAACATAGGAGAACTTAATCTTAATACCGATTAATCTTAAGAGAGAAACAGGAGATC 203
Qy      177  GAAAGACCTTCCGCTTTTGGAGCGCGGATGCTGATGATGATGATGATGATGATGATG 236
Db      204  GCAAGACCTTCCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 263
Qy      237  CTACCAAGCGCAGATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 296
Db      264  CCACCAAGCGCAGATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 323
Qy      297  CACGCGCCAGATCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 356
Db      324  CACGCGCCAGATCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 383
Qy      357  GATCCAGATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 416
Db      384  GATCCAGATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 443
Qy      417  AAGAAAGGTTAAGGTTAATATGCTGATGATGATGATGATGATGATGATGATGATGATG 476
Db      444  AAGAAAGGTTAAGGTTAATATGCTGATGATGATGATGATGATGATGATGATGATGATG 503
Qy      477  TAACTACGTGCGCAGAGCGCGGTAATAGTAAAGGTTGCAAGGCTTAATGGAATTA 536
Db      504  TAACTACGTGCGCAGAGCGCGGTAATAGTAAAGGTTGCAAGGCTTAATGGAATTA 563
Qy      537  GCGTAAAGGTTGCGCAGAGCGCGCTTGTATGATGATGATGATGATGATGATGATGATG 596
Db      564  GCGTAAAGGTTGCGCAGAGCGCGCTTGTATGATGATGATGATGATGATGATGATGATG 623
Qy      597  GAATTCGCTTAAACTCAAGGCTTAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 656
Db      624  GAATTCGCTTAAACTCAAGGCTTAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 683
Qy      657  GCAATGAAATGCGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 716
Db      684  GCAATGAAATGCGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 743
Qy      717  CACTGACGTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 776
Db      744  CACTGACGTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 803
Qy      777  CGCCCTAAACGATGCACTAATGTTGGGCTTATTAAG--CTTGGTAAACGAGTAA 834
Db      804  CGCCCTAAACGATGCACTAATGTTGGGCTTATTAAG--CTTGGTAAACGAGTAA 863
Qy      835  GCGTAAAGTGAACCGCTGGGAGTACGCTGCAAGATTAACTCAAGAAATTGACG 894
Db      864  GCGTAAAGTGAACCGCTGGGAGTACGCTGCAAGATTAACTCAAGAAATTGACG 923
Qy      895  GGAACCGGCAAGAGCGTGAATTAATGATTAATTCGATGCAACGGGAAACCTTA 954
Db      924  GGGCCCGCAAGAGCGTGAATTAATGATTAATTCGATGCAACGGGAAACCTTA 983
Qy      955  ACCCTGACATGAGGAATTTTCTAGAGATGATTAGTGG-CTTGGGAAACGTTACACA 1013

```



Db	984	GGCCTTGACATGCACGGAACTTTCCAGAGATGGATTGGTGCTTCGGGAAACCGTGACACA	1043
Qy	1014	GGTGTGCATGGCTGTGCTGCAGCTTGCTGTGCGTAGATGTTGGGTAAAGTCCCGCAACGAG	1073
Db	1044	GGTGTGCATGGCTGTGCTGCAGCTGTGTCTGGAGATGTTGGGTAAAGTCCCGCAACGAG	1103
Qy	1074	CGCAACCCCTTGCTAATTAAATGGCAATCA--TTTGGTTGGGACCTTAAATAGACCTGGCCGGT	1131
Db	1104	CGCAACCCCTTGCTAATTGGTCCACAGTAATGTGTGGAACTCTAAGGAGACCGCCGGC	1153
Qy	1132	GACAAACCGAGGAAAGGTGGGATGACGTCAAGTCCATGCGCCCTTAATGGGTAGGGCTT	1191
Db	1164	GACAAAGCCGAGGAAAGTGGGGATGACGTCAAGTCATGATGACCTTACGCGCACGGGGCTA	1223
Qy	1192	CACACGTAAATCAATGCGGCGGTACACAGAGGTTCACACCCCGAGAGGGAGACTAAATCTCA	1251
Db	1224	CACACGTAACTCAAAATGTGGGGACAGAGGGCTGCAAACTCCGGAAGTGAACCAATCCCA	1283
Qy	1252	GAAAGCGAGTGTAGTCCGGAATCGAGATCTGCACATCTCCGTGAAGTGGGAATCGCT	1311
Db	1284	GAAACCCCATCTCACTCCGGAATGGAGTCTGCAATCTGCACTGCATCTCAAGAGTGGAAATCGCT	1343
Qy	1312	AGTAATCGCGGATCAGCA-TGTCTCGGTGAATAGTTCCTCCGGGTTTGTACACACCGGCC	1370
Db	1344	AGTAATCCCAATCAGCAATGTCTGGGTGAATACGTTCCCGGCTTTGTAAACACACCGGCC	1403
Qy	1371	GTCACACCATGGAGATGGGATTTCAACAGAGAGAGGTAGTCTTAACCGTAAGAGGGCGCTT	1430
Db	1404	GTCACACCATGGAGATTTGTTCGACACAGAGAGAGGTAGTCTTAACCGCAAGGGGGCGCTT	1463
Qy	1431	GCCACGGTGAATTCATGACTGGGGTG	1457
Db	1464	GCCACGGTGTGGCCCATGACTGGGGTG	1490

```

RESULT 16
US-09-967-376-1
; Sequence 1, Application US/09967376
; Patent No. 6482635
; GENERAL INFORMATION:
; APPLICANT: WALKER, Harrell L.
; APPLICANT: HIGGINBOTHAM, Lawrence R.
; TITLE OF INVENTION: CONTROL OF CYANOBACTERIA WITH A BACTERIUM
; FILE REFERENCE: 013243-0007
; CURRENT APPLICATION NUMBER: US/09/967,376
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1540
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Bacterium SG-3; gram negative;
; OTHER INFORMATION: rod-shaped; exhibits flagellate motility; pathogenic to
; OTHER INFORMATION: cyanobacteria and algae; yellow colonies on BG-11 medium suppl.
; US-09-967-376-1

```

Query Match	72.3%	Score 1053	DB 3	Length 1540
Best Local Similarity	84.7%	Pred. No. 0		
Matches 1242	Conservative	0	Mismatches 215	Indels 10
			Gaps	5
QY	1	ATTGAACCTGGCGGCATGCTTTACATCATGCAAGTCGAA	CGGCAGACGAGATGCTTGAT	60
DB	24	AGTGAACCTGGCGGCAGGCTTAACATCATGCAAGTCGAA	CGGCAGACAGCATGTCAT	83
QY	61	CT---GGTGGCGAGTGGCGGACGGGTGATTAATGCATGGA	ACGTATCCAGAAAGGGG	116
DB	84	ACTGTGGGTGGGAGTGGCGGACGGGTGAGGAATTCATCGA	AACTTGCCCACTCGTGGGG	143
QY	117	GGTAAAGCATCGAAAGATGTGCTAATACCGCATATATCTTA	AGAGGAAAGCAGGGGATC	176

Db	144	GATAACATAGGAAACCTTATCTCTAATACCCGATACGACTGACGGTGTAAACGGGGGATC	203
Qy	177	GAAGACCTTGCCGCTTTTGGAGCGCGCGATGCTGATTTAGCTAGTTGGTGGGTAAAGCC	236
Db	204	GCAAGACCTCGCGCGATTGGATGAGCCGATGCTCGATTAGCTAGTTGGCGGGTAAATGGC	263
Qy	237	CTACCAAGCGACGATCGATGATTGTTGTCGAGAGGACGACGACGACACTGGGACTGAGA	296
Db	264	CCACCAAGCGACGATCGGTTAGCTGTTGTGAGAGGATGATCAGCCACCTGGAACTGAGA	323
Qy	297	CACGGCCGAGACTCTCTACGAGAGGACGACGTCGGGAAATTTTGGACAAATGGCGCAAGCT	356
Db	324	CACGGTCCAGACTCTTACGGGAGGACGACGATGGGGAATATTGACAAATGGGCGCAAGCT	383
Qy	357	GATCCAGCAATGCCCGTAGTGAAGAGGCTTTCCGGTGTAAAGCTCTTTCAGTCSAG	416
Db	384	GATCCAGCATGCCCGGTGGGTGAAGAGGCTTTCCGGTGTAAAGCCTTTTGTGTCGGA	443
Qy	417	AAAGAAAGGTTACGGTAAATTAATGTGTACTATGACGCTATCGACAGAGAAAGCACCCGGC	476
Db	444	AAAGAAAGCATCGATTAATTAATCTTCGCGTGTCTGACGGTACCGGAAGAAATAGCACCCGAC	503
Qy	477	TAACTAGGTGCAGAGCCGCGGTAAATCTAGAGGGTGCAGCGTTAATCCGAAATTAATCGG	536
Db	504	TAACCTTGTCGACAGACCCGCGTAAATCGAAGGGTGCAGCGTTAATCTCGAAATTAATCGG	563
Qy	537	GCGTAAAGGATGCGCAGCGCGCTTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTCGG	596
Db	564	GCGTAAAGCGTAGGTTGGTTGTTAAAGCTGATGTGAAGACCTGGGCTCAACCTGG	623
Qy	597	GAATTTGCTTTGAATCTCAAGGCTAGAGTGTGCGACAGGAGGAGTGGAAATTCATGTGCTA	656
Db	624	GAATTTGCAATTGGAATCTCGCGCGGCTAGAGTGCGGTAGAGATGCGGAAATTCGCGGTGA	683
Qy	657	GCAGTGAATCGTAGAGATATGGAAGAACTGCATGGCGAAGCGAGCTCTGGGTAAAT	716
Db	684	GCAGTGAATCGTAGAGATCGGAGAGAACACTCTGTGCGAAGCGCGCACTTGGAACAG	743
Qy	717	CACTGACGCTATGACCGAAAGCGTGGGGACAACAGAGTTAGTAACCTGTAGTGTCA	776
Db	744	CACTGACACTGAGGAGCGAAAGCGTGGGGACAACAGAGTTAGTAACCTGTAGTGTCA	803
Qy	777	CGCCCTAAACGATGTCAACTAGTTGTTGGGCTTATTAGG--CTTGGTAAAGAACTTAAC	834
Db	804	CGCCCTAAACGATGTCAACTGAGTTGGGTGCAACTAGGCACTCAGATATGAAGCTTAAC	863
Qy	835	GCGTGAAGTTTACCGGCTGGGGGAGTACGCGTGCAGAAATTAAACCTCAAGGAATTGACGG	894
Db	864	GCGTTAAGTTGCGCGCTGGGGGAGTACGCGTGCAGAACTCAAGGAATTGACGG	923
Qy	895	GGACCCGACAAAGCGGTGAGATTATGTGATTAAATTCGATGCGACGCGAAAACTTAACCT	954
Db	924	GGGCGCGACAAAGCGGTGAGATTATGTGTTAAATTCGATGCGACGCGAAAACTTAACCT	983
Qy	955	AACCTTGACATGTAGCGAATTTTCTAGAGATGATTAAGTG--CTTGCGGAACGCTAACACA	1013
Db	984	GGCCTTGACATGACGGAACCTTCCAGAGATGAGATTGGTGCCTTCGCGGAACCGTGACACA	1043
Qy	1014	GGTGCTGATAGGCTGTGCTCAGGCTTCGTGCTGAGATGTTGGGTTAACTCCCGAACGAG	1073
Db	1044	GGTGCTGATAGGCTGTGCTCAGGCTTCGTGCTGAGATGTTGGGTTAACTCCCGAACGAG	1103
Qy	1074	CGCAACCTTGATCAATTAATGSCATCA--TTTGGTGGGCACTTAAATGAACTGCGCGT	1133
Db	1104	CGCAACCTTGATCCCTAGTTGCTCAGACGCTAAATGTTGGAACTCTAGGGAACCGCGCGC	1166
Qy	1132	GACAAACCGAGGAAGGTGGGATGACGTCAAGTCTTCATGAGCCCTTAATGGGTAGGCTT	1191
Db	1164	GACAAAGCGGAGGAAGGTGGGATGACGTCAAGTCAATCATGAGCCCTTAATGGGACGAGCTA	1222
Qy	1192	CACACGTAAATCAATGCGCGCTACAGAGGTTTGCACACCGCGAGAGGGGACCTAATCTCA	1251
Db	1224	CACACGTAAATCAATGCGCGGGAACAGAGGGCTGCAAACTCCGAGAGTGAACCAATCCCA	1283





```
RESULT 18
US-09-735-567-3
; Sequence 3, Application US/09735567
; Patent No. 6608190
; GENERAL INFORMATION:
; APPLICANT: Brannucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE REFERENCE: BCI033 US NA
; CURRENT APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Denitrifying bacteria
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RA1
US-09-735-567-3
```

```
Query Match      70.4%; Score 1026.4; DB 3; Length 1450;
Best Local Similarity 84.4%; Pred. No. 0; Mismatches 211; Indels 14; Gaps 5;
Matches 1217; Conservative 0;
```

```
QY 1 ATTGAACGCTGGCGGCGATCTTTACATGACATGCAAGCGGACGCGATGCTTGCAT 60
DB 3 ATTGAACGCTGGCGGCGATCTTTACATGACATGCAAGCGGACGCGATGCTTGCAT 58
QY 61 CTGGTGGCGAGTGGCGGACGCGGTGAATGCAATCGGAACGTAATCCAGAAAGGGGGGTA 120
DB 59 CTGGGCGGCGAGCGGCGAAGCGGTGAATGCAATCGGAACGTAATCCAGAAAGGGGGGTA 118
QY 121 ACGCATCGAAGAAATGCTAATACCGCATATCTCTAAGAGAGAAAGCGGGGATCGAA 180
DB 119 GCCCGGCGAAGCGGATTAATACCGCATATCTCTAAGAGAGAAAGGGGGGATCGAA 178
QY 181 GACCTTGGCGCTTTGGAGCGCGCATGCTGATTAAGTAGTGGTGGGTAAGGCTTAC 240
DB 179 GGCCTCAGCGCTTTGGAGCGCGCATGCTGATTAAGTAGTGGTGGGTAAGGCTTAC 238
QY 241 CAAGCGAGCATCAAGTAGTGGTCTGAGAGAGACGACACCTGGGATCGAGACAG 300
DB 239 CAAGCTGCGATCTGATGCTGGTCTGAGAGAGATGATCAGCCACACTGGGATCGAGACAG 298
QY 301 GCCCGAGCTCTTACCGGAGGCGAGCGATGGGGAATTTGGCAATGGGCGCAAGCCTGATC 360
DB 299 GCCCGAGCTCTTACCGGAGGCGAGCGATGGGGAATTTGGCAATGGGCGCAAGCCTGATC 358
QY 361 CAGCAATGCGCGGTGAGTGAAGAGCGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAGA 420
DB 359 CAGCAATTCGGGTGAGTGAAGAGCGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAGA 418
QY 421 AAAGGTTACGGTAATATCTGATCATGACGATTCGACAGAAAGACCGGCTTAC 480
DB 419 AAAGGCTCTCTTAATACAGGGGCGCATATGACGATTCCTTAAGAAATTAAGCACCGGCTTAC 478
QY 481 TACCGTCCAGCAGCGCGCGGTATATGATGAGGTCGAAGCGTTAATGGAAATTACTGGGCGT 540
DB 479 TACCGTCCAGCAGCGCGCGGTATATGATGAGGTCGAAGCGTTAATGGAAATTACTGGGCGT 538
QY 541 AAAGGTTGCGAGCGCGCTTTGTAAGTCAAGTGTAAATCCCGGGCTTAACTGGGAAT 600
DB 539 AAAGGTTGCGAGCGCGCTTTGTAAGTCAAGTGTAAATCCCGGGCTTAACTGGGAAT 598
QY 601 TGGCTTTGAACTACAAAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGACAG 660
DB 599 TGGCTTTGTAAGTCAAGGCTGAGTGTGCGAGAGGAGTGAATTCATGTGACAG 658
```

```
QY 661 TGAATGCGTAGAGATATGAAGAACATCGATGCGAAGCGAGCGCTCTGGGTTAACACT 720
DB 659 TGAATGCGTAGAGATATGCGAGAGAACATCGATGCGAAGCGAGCGCTCTGGGTTAACACT 718
QY 721 GACGCTCAATGACGAAACCGTGGGAGAGCAACAGATTAATGATACCTGTGATCCAGCGC 780
DB 719 GACGCTCAATGACGAAACCGTGGGAGAGCAACAGATTAATGATACCTGTGATCCAGCGC 778
QY 781 CT-AAACGATGCAACTGTTGTTGGGCTTATTAAGCTTGTAAAGGATTAAGCGGCGT 839
DB 779 CTAAACGATGCAACTGTTGTTGGGATTAATTAATGATGCAACGCGAAGGATTAAGCGGCGT 838
QY 840 AAGTTACCGCTGGGAGAGTACGCTGCGAAGATTAACTAAAGAAATTGACGCGGAGC 899
DB 839 AAGTTACCGCTGGGAGAGTACGCTGCGAAGATTAACTAAAGAAATTGACGCGGAGC 898
QY 900 CGCACAGCGGTGATTAATGATGCAATTAATGATGCAACGCGAAGGATTAAGCGGCGT 959
DB 899 CGCACAGCGGTGATTAATGATGCAATTAATGATGCAACGCGAAGGATTAAGCGGCGT 958
QY 960 TGAATGCGTAGAGATATGCGAGAGAACATCGATGCGAAGCGAGCGCTCTGGGTTAACACT 1016
DB 959 TGAATGCGTAGAGATATGCGAGAGAACATCGATGCGAAGCGAGCGCTCTGGGTTAACACT 1018
QY 1017 GCTGATGCGTGTGCTGAGTCTGCTGAGATGTTGGGTTAAGTCCCGCAAGAGCGC 1076
DB 1019 GCTGATGCGTGTGCTGAGTCTGCTGAGATGTTGGGTTAAGTCCCGCAAGAGCGC 1078
QY 1077 AACCTTGTCAATTAATGCTCAATTTGGTGGGCACTTTAATGAGTCTGCGGTGACAA 1136
DB 1079 AACCTTGTCAATTAATGCTCAATTTGGTGGGCACTTTAATGAGTCTGCGGTGACAA 1133
QY 1137 ACCGAGAGAGTGGGAGTGAAGTCAAGTCTCAATGCGCTTAATGGGTTAAGGCGCTTCAAC 1196
DB 1134 ACCGAGAGAGTGGGAGTGAAGTCAAGTCTCAATGCGCTTAATGGGTTAAGGCGCTTCAAC 1193
QY 1197 GTAATACATAGGCGGTCACAGAGGTTGCAACCGCGAGGAGGAGTCAATCTGCAAGAG 1256
DB 1194 GTCAATACATAGGCGGTCACAGAGGTTGCAACCGCGAGGAGGAGTCAATCTGCAAGAG 1253
QY 1257 GCGGTGTAATGCGGATGCGAGTCTGCACTCGGTGAAGTGGGATCGCTTAA 1316
DB 1254 CCAATGCTGATGCGGATGCGAGTCTGCACTCGGTGAAGTGGGATCGCTTAA 1313
QY 1317 TCGGATGATCAGATGTCGCGGTGAATACGTT-CCCGGCTCTTGAACACCGCGCTGAC 1375
DB 1314 TCGGATGATCAGATGTCGCGGTGAATACGTT-CCCGGCTCTTGAACACCGCGCTGAC 1373
QY 1376 ACCATGGAAGTGGGTTTCAACGAGAGAGTGAATCTAACCCTAAGAGAGGCGCTTGCAC 1435
DB 1374 ACACCATGAGAGGTTTCCGCAAGAGTGAATAGCTTAACCGCAAGAGGCGGATTAACAC 1433
QY 1436 GG 1437
DB 1434 GG 1435
```

```
RESULT 19
US-09-793-920A-1
; Sequence 1, Application US/09793920A
; Patent No. 6479621
; GENERAL INFORMATION:
; APPLICANT: Canon Inc.
; TITLE OF INVENTION: Polyhydroxyalkanoate containing 3-hydroxyphenylalkanoic acid as
; FILE REFERENCE: monomer unit, and method for producing the same.
; CURRENT APPLICATION NUMBER: US/09/793,920A
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii 161 strain.
```

US-09-793-920A-1

Query Match 70.2%; Score 1022.6; DB 3; Length 1501;  
Best Local Similarity 83.6%; Pred. No. 0;  
Matches 1219; Conservative 0; Mismatches 234; Indels 6; Gaps 5;

```
QY      3 TGAACGCTGGCGCATGCTTTTACACATGCAAGTCGAAACGGACACGATGCTTGATCT 62
DB      1 TGAACGCTGGCGCGACCTTAACACATGCAAGTCGAAACGGACACGATGCTTGATCT 59
QY      63 GGTGCGAATGGCGGACCGGCTGATGATGATGCAAGCTATCCAGAAAGGGGGGTAAAC 122
DB      60 GAATTC-ACGGCGGACGGGTGATGATGCTGATGATGCTGCTGGTATGGGGAACAC 118
QY      123 GCATCGAAAGATGTGCTTAATACCGCATATCTTAAGAGAAAGCAAGCGGATCGAAAGA 182
DB      119 GTCTCGAAAGGACGCTTAATACCGCATATGCTCTACGGAGAAAGACGGGACCTTCGGG 178
QY      183 CCTTGCGCTTTTGAAGCGGCGCATGCTGATTAAGTTAGTTGGGTAAAGGCTTACCA 242
DB      179 CCTTGCGCTATCAGATGAGCTTATAGTCGATTAAGTTGGGTGATGATGCTTACCA 238
QY      243 AGCGCAGATCAGTATGCTGCTGAGAGAGACCAACCACTGGGATCGAGACACGCGC 302
DB      233 AGCGCAGATCAGTATGCTGCTGAGAGAGATGATCAGTCACTGGAACCTGAGACACGCT 298
QY      303 CCAGACTCTCAACGGAGGAGGAGAGAGGAAATTTTGAACATGGGCGCAAGCTGATCCA 362
DB      299 CCAGACTCTCAACGGAGGAGGAGAGAGGAAATTTTGAACATGGGCGCAAGCTGATCCA 358
QY      363 GCAATGCCCGCTGATGAGAAAGGCTTCGGGTTGTAAAGCTTTTCACTGAGAAAGAA 422
DB      359 GCCATGCCCGCTGATGAGAAAGGCTTCGGGTTGTAAAGCTTTTCACTGAGAAAGAA 418
QY      423 AGCTTACGCTTAATATCGTGAATCATGACGCTATGACAGAAAGACCGGCTTAACCTA 482
DB      419 GGCATTTACCTTAATATCGTGAATCATGACGCTATGACAGAAAGACCGGCTTAACCTA 478
QY      483 CGTGCAGAGCGCGGCTAATACGTAAGGCTGCAAGGCTTAATCGAATTAATCTGGGCTAA 542
DB      479 TGTGCAGAGCGCGGCTAATACGTAAGGCTGCAAGGCTTAATCGAATTAATCTGGGCTAA 538
QY      543 AGGCTGCGAGCGGCTTTGTAAAGTCAATGATGAAATCCCGGGCTTAACTCGGAATTTG 602
DB      539 AGGCTGCGAGCGGCTTTGTAAAGTCAATGATGAAATCCCGGGCTTAACTCGGAATTTG 598
QY      603 CGTTTGAACCTACAGGCTTATGATGCTGCGAGAGGAGGTGGAATTCATGTGTAGCAAGT 662
DB      599 CATTCAAACTGACAGCTATGATGATGATGAGAGGTGATGAAATTTCTGTGTAGCGGTG 658
QY      663 AAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
DB      659 AAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
QY      723 CGCTCAGTACAGAAAGCTGCGGAGGAGCAACAGATTAAGATACCTGTGATGCTACCGCTT 782
DB      719 CACTGAGGTGCGAAAGGTGCGGAGGAGCAACAGATTAAGATACCTGTGATGCTACCGCTT 778
QY      783 AAACGATGCTACCTAGTTGTTGGGCTTTATTAAG- GCTTGTGTAACGAAGCTTAACGGCTGAA 841
DB      779 AAACGATGCTACCTAGTTGTTGGGAGGCTTTAGCTTCTAGTGGCGCAAGCTTAACGGCTTAA 838
QY      842 GTTGAACCGGCTGGGAGATACGCTGCGAAAGATTAACCTCAAAAGAAATTTGACGGGAGACCG 901
DB      839 GTTGAACCGGCTGGGAGATACGCGCGCAAGGTTAAATCTCAAAATGAATTTGACGGGAGACCG 898
QY      902 CACAAGCGGTGATGATGATGATGATTAATTCAGATGCAACGCAAAACCTTACCTACCTTGG 961
DB      899 CACAAGCGGTGATGATGATGATGATTAATTCAGATGCAACGCAAAACCTTACCTACCTTGG 958
QY      962 ACATGTAAGCAATTTTCTAAGATAGTATGTTG- CTTTGGGAAACGCTTAACACAGGTGCTG 1020
DB      959 ACATGTAAGCAATTTTCTAAGATAGTATGTTGTTGCTTGGGAAACATTTGAACAGGTGCTG 1018
```

```
QY      1021 CATGCTGTGCTCAGCTGCTGCTGAGATGTTGGGTTAACTCCCGCAACGAGCGCAACC 1080
DB      1019 CATGCTGTGCTCAGCTGCTGCTGAGATGTTGGGTTAACTCCCGTAACGAGCGCAACC 1078
QY      1081 CTGTGCTTAATATGCTCATCATTTT- GCTTGGGCACTTAAATAGACCTCCGCTGACAAAC 1138
DB      1079 CTGTGCTTAATATGCTCATCATTTT- GCTTGGGCACTTAAATAGACCTCCGCTGACAAAC 1138
QY      1139 CGAGAGAGGTGGGATGACGTCAGATGCTCAATGAGCTTATGAGTGGCTTACACAGT 1198
DB      1139 CGAGAGAGGTGGGATGACGTCAGATGCTCAATGAGCTTATGAGTGGCTTACACAGT 1198
QY      1199 AATTAATGAGCGCGCTGACAGAGGTGGCCAAACCCCGAGAGGGAGGATTAATTCAGAAAGG 1258
DB      1199 GCTTACAAATGCTGCTGACAGAGGTGGCCAAACCCCGAGAGGTGAGCTTAATTCAGAAAC 1258
QY      1259 CGTGTGATCGGAGATCGAGATGCTGCAACTCGACTCGTGAAGTGGAAATCGCTAATATC 1318
DB      1259 GATGTGATCGGAGATCGAGATGCTGCAACTCGACTCGTGAAGTGGAAATCGCTAATATC 1318
QY      1319 GCGGATCAGCATGCTGCGGTGATACGTTCCCGGCTCTTGTACACACCGCCGTCACACC 1378
DB      1319 GCGGATCAGCATGCTGCGGTGATACGTTCCCGGCTCTTGTACACACCGCCGTCACACC 1378
QY      1379 ATGGAGTGGGTTTACCAAGAAAGCATGTAATCTTAACCTGAAGAGGGCGCTTGCACGCT 1438
DB      1379 ATGGAGTGGGTTTACCAAGAAAGCATGTAATCTTAACCTGAAGAGGGCGCTTGCACGCT 1438
QY      1439 GAGATTCATGATGGGGTG 1457
DB      1439 GTGATTCATGATGGGGTG 1457
```

RESULT 20

```
US-09-821-016-5
; Sequence 5, Application US/09821016
; Patent No. 6485951
; GENERAL INFORMATION:
; APPLICANT: CANON INC.
; TITLE OF INVENTION: Polylhydroxalkanoate Synthase and Gene Encoding the Same Enzyme
; FILE REFERENCE: 4051021
; CURRENT APPLICATION NUMBER: US/09/821,016
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jesseni P161 ; BP-7376
; FEATURE:
;
US-09-821-016-5
```

Query Match 70.2%; Score 1022.6; DB 3; Length 1501;  
Best Local Similarity 83.6%; Pred. No. 0;  
Matches 1219; Conservative 0; Mismatches 234; Indels 6; Gaps 5;

```
QY      3 TGAACGCTGGCGCATGCTTTTACACATGCAAGTCGAAACGGACACGATGCTTGATCT 62
DB      1 TGAACGCTGGCGCGACCTTAACACATGCAAGTCGAAACGGACACGATGCTTGATCT 59
QY      63 GGTGCGAATGGCGGACCGGCTGATGATGATGCAAGCTATCCAGAAAGGGGGGTAAAC 122
DB      60 GAATTC-ACGGCGGACGGGTGATGATGCTGATGATGCTGCTGGTATGGGGAACAC 118
QY      123 GCATCGAAAGATGTGCTTAATACCGCATATCTTAAGAGAAAGCAAGCGGATCGAAAGA 182
DB      119 GTCTCGAAAGGACGCTTAATACCGCATATGCTCTACGGAGAAAGACGGGACCTTCGGG 178
QY      183 CCTTGCGCTTTTGAAGCGGCGCATGCTGATTAAGTTGGTGGGCTTAAAGGCTTACCA 242
DB      179 CCTTGCGCTATCAGATGAGCTTATAGTGAATTAAGTTGGTGAAGTGAATGCTTACCA 238
```

OY	243	AGGCGACGATCAGTAGTTGGTCTGTGAGAGACGACCAGCACTGTGGACTGAGACGGC	302
Db	239	AGGCGACGATCCGTGAATCTGTCTGTGAGAGATGATCAGTCACTGGAACTGAGACCGT	298
OY	303	CCAGACTCTCTACGGGAGGACGAGTGGGGAAATTTTGGACAATGGGCGCAAGCTGTATCA	362
Db	299	CCAGACTCTCTACGGGAGGACGAGTGGGGAAATTTTGGACAATGGGCGAAAGCTGTATCA	358
OY	363	GCAATGCCCGCTGAGTGAAGAAAGCCTTCCGGTGTGTAAAGCTCTTTCAGTGGAGAA	422
Db	359	GCAATGCCCGCTGTGTGAAGAAAGCTTCCGATGTGTAAAGCACTTAAAGTGTGGAGAA	418
OY	423	AGGTTACGGTAATATATGTGCTCATGACGGTATCGACAGAGAAAGACCCGGTAACTA	482
Db	419	GGCATTTAACTTAATATCGTTAGTGTTTGACGTTTACCGACAGAAATGACCGGGCTAATC	478
OY	483	CGTGCCACACCGCGGGTAAATCGTAGGGTGCAGACGGTAATCCGAATTACCTGGGCGTAA	542
Db	479	TGTGCTACACAGCCGGGTAAATCAGAGGGTGCAGACGGTAATCCGAATTACCTGGGCGTAA	538
OY	543	AGGATGCCGACGGCGGCTTTGTAACTGTCAATGTGAATATCCCGGGCTTAACTGGGAAATG	602
Db	539	AGCGCGGTAAGTGGTGTTTGTAACTGTGAATGTGAAGACCCGGGCTCAACTGGGAAATG	598
OY	603	CGTTTGAACATAAGGCTAGAGTGTGCGAGGAGGGTGCAGATTCATGTGTAGAGCTG	662
Db	599	CATTCAAAATCTGACAAAGCTAGAGTGTGTAGAGGGTGTGTGAATTTCTGTGTAGCGGTG	658
OY	663	AAATGCGTAGAGATATGGAAGAACATCGATGGCGAAAGGACGCTCTGGGTAAACATCTGA	722
Db	659	AAATGCGTAGATATAGAAAGGAACACAGTGGCGAAAGGACACCTTGAATCTGATCTGAA	718
OY	723	CGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTAGTCCAGCCCT	782
Db	719	CACGTAGGTGCGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTAGTCCAGCCGT	778
OY	783	AAACGATGTCAACTAGTGTGTGGCCCTTATTAG-GCTTGTGTACGAACTAACCGGTGA	841
Db	779	AAACGATGTCAACTAGCGTGTGGAGCCCTTGTAGCTCTTAATGTGGCCGACCTAACCGATTAA	838
OY	842	GTTGACCGCGTGGGAGATACCGTGCACAGATTAAACTCAAGAAATTGACGGGGACCG	901
Db	839	GTTGACCGCGTGGGAGATACCGCGCCGACAGTTAAACTCAATGTAAATTGACGGGGACCG	898
OY	902	CACAAGCGGTGATTAATGTGATTAATTCGATGCAACCGCAAAAACCTTAACTCACTCTTG	961
Db	899	CACAAGCGGTGAGCATGTGTATTATTCGAAAGCAACGGAAGAACCTTAACTCAAGCCCTTG	958
OY	962	ACATGTAGCGAAATTTTCTAGAGATAGATTATAGT-CTTCGGGAAACGCTAACACAGGTGCTG	1020
Db	959	ACATCCAAATGAATTTTCCAGAGATGATGATGGGTGCTTCGGGAAACATTGACAGGTGCTG	1018
OY	1021	CATGGCTGTCTCACTCGTGTCTGTGAGATGTTGGGTTAAATGCCCGCAACGACGGCAAC	1080
Db	1019	CATGGCTGTCTCACTCGTGTCTGTGAGATGTTGGGTTAAATGCCCGTAAACGACGGCAAC	1078
OY	1081	CTTGTCAATTAATTTGCAATCTT--GGTTGGGCACTTAAATGAGACTCCGGTGAACAAC	1138
Db	1079	CTTGTCTCTTAAGTTACGACGATTAATGTGTGGCACTCTTAAGAGACTCCGGTGAACAAC	1138
OY	1139	CGAGAAAGTGGGAGATGACGTCAAGTCTCTCATGGCCCTTAATGGGTAGGGCTTCAACGT	1198
Db	1139	CGAGAAAGTGGGAGATGACGTCAAGTCTCTCATGGCCCTTAATGGGTAGGGCTTCAACGT	1198
OY	1199	AATACATAGCGCGGTACAGAGGGTTGCCAACCCCGAGAGGGAGCTAATCTCAGAAAGCG	1258
Db	1199	GCTACAAATGTGTGGTACAGAGGGTTGCCAACCCCGAGAGGGAGCTAATCTCAGAAAGCG	1258
OY	1259	CGTGTAGTCCGGATCCGAGTCTGCAACTCGACTCCGTGAAGTCCGAATCCGTAGTATCT	1318
Db	1259	GATCGTAGTCCGGATCCGAGTCTGCAACTCGACTCCGTGAAGTCCGAATCCGTAGTATCT	1318
OY	1319	GGGATGACAGCATGTGCGGGTGAATAGTTCCCGGGTCTTTGTACACACCGCCCTGCACAC	1378

Query	Match	70.2%	Score 1022.6	DB 3	Length 1501			
Query	Beet Local Similarity	83.6%	Pred. No. 0	Mismatches 234	Indels 6	Gaps 5		
Matches 1219	Conservative	0	Mismatches	234	Indels	6	Gaps	5
Db	1319	GCAGATCGAATGTCGCGGTAAATACGTTCCCGGGCTTTGACACACCGCCCGTCAACC	1378					
Qy	1379	ATGGAGTGGGTTTCACCAAGACAGGTAGTCTAACCGTAAGAGGGCGCTTGGCACCCT	1438					
Db	1379	ATGGAGTGGGTTGCACCAAGATAGTACTAACCTTCGGAGAGACGGTTACACCGT	1438					
Qy	1439	GAGATTCACTGCTGGGCTG	1457					
Db	1439	GTGATTCACTGACCTGGGCTG	1457					
RESULT 21								
US-09-745-476-1								
Sequence 1, Application US/09745476								
Patent No. 6521429								
GENERAL INFORMATION:								
APPLICANT: CANON INC.								
TITLE OF INVENTION: Preparation of poly-hydroxyalkanoic Acid								
FILE REFERENCE: 4351008								
CURRENT APPLICATION NUMBER: US/09/745,476								
CURRENT FILING DATE: 2000-12-26								
NUMBER OF SEQ ID NOS: 1								
SOFTWARE: Microsoft Word								
SEQ ID NO 1								
LENGTH: 1501								
TYPE: DNA								
ORGANISM: Pseudomonas jessenii Pl61 ; FERM P-17445								
US-09-745-476-1								
Query Match	70.2%	Score 1022.6	DB 3	Length 1501				
Beet Local Similarity	83.6%	Pred. No. 0	Mismatches 234	Indels 6	Gaps 5			
Matches 1219	Conservative	0	Mismatches	234	Indels	6	Gaps	5
Qy	3	TGAACGCTGGGCGGATCTTTACACATGTCGAAACGACGACGACGAGTCTTGATCT	62					
Db	1	TGAACGCTGGGCGGACGCTTACACATGTCGAAACGAGCGAGTACCGGAGCTTCTCT	59					
Qy	63	GGTGGCAGTGGCGGACGGGTGATGATGATCGAACGTATCCAGAAAGAGGGGGTAACT	122					
Db	60	GAAATTC-AGCGGCGGACGGGTGATGATGATGCTAGGAATCTCTGTTAGTGGGGGCAAC	118					
Qy	123	GCATCGAAGATGCTTAATACCGATTAATCTTAAGGAGGAAAGACAGGGGATCGAAGA	182					
Db	119	GTCTCGAAGGAGGACCTTAATACCGATTAATGCTTAACGGAAGAAAGCAAGGACCTTGGGG	178					
Qy	183	CCCTGCGCTTTTGGAGCGGCGGATGCTGATTAAGTAAAGTCTTTCAGTGGGCTTACCA	242					
Db	179	CCCTGCGCTTACAGATGAGCTTAAAGTGGGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAG	238					
Qy	243	AGCGACATGATGATGTTGTTCTGAGAGGACGACCGCACACTGGGACTGAGACAGGC	302					
Db	239	AGCGACATGATGATGTTGTTCTGAGAGGATGATGATGATGATGATGATGATGATGATGAT	298					
Qy	303	CCAGACTCTTACGAGAGGACGATGAGGAAATTTTGGACATATGGGCGGAAGCTTGAATCA	362					
Db	299	CCAGACTCTTACGAGAGGACGATGAGGAAATTTTGGACATATGGGCGGAAGCTTGAATCA	358					
Qy	363	GCAATGCGGCGGTGTAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGGGAGGAA	422					
Db	359	GCAATGCGGCGGTGTAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGGGAGGAA	418					
Qy	423	AGTTTACGTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	482					
Db	419	GCAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	478					
Qy	483	CGTGCACGACGCGCGGTAAATCGTAAAGGTGACAGCTTAAATCGAATTAATCGGGCTTAA	542					
Db	479	TGTGTCACACGACGCGCGGTAAATCGAAGGTGACAGCTTAAATCGAATTAATCGGGCTTAA	538					
Qy	543	AGGGTGGCAGGCGGCTTGTAAAGTCAATGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA	602					
Db	539	AGGGTGGCAGGCGGCTTGTAAAGTCAATGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA	598					

QY	603	CCTTGGAAACTCAACGGCTAGAGTGTGGCAGAGGGAGGTGGAAATTCATGTGTAGCACTG	662
Db	599	CATTCAAAAACCTGACCAAGCTAGAGATATGTATAGAGGGTGTGGAAATTTCTGTGTACGGGTG	658
QY	663	AAATCGGTAGAGATATAGGAAGAAATATCGATGCGCAGAGCGACGCTCTGGGTTAACTCTGA	722
Db	659	AAATCGGTATATATAGGAAGGAACACCATGTGGCGAAGGCCACACTGGAATGATATCTGA	718
QY	723	CGCTCATGCAACGAAGCGTGGGGAGCAACAGATATAGATACCTCTGTATGTCCAGCCCT	782
Db	719	CACCTGGGGTGGCGAAGACGTGGGGAGCAAAACAGATATAGATACCTCTGTATGTCCAGCCCT	778
QY	783	AAACGATGTCACTAGTGTGTGGGCTTATTTAG--GCTTGGTAAACGAACTTAAACCGGTGA	841
Db	779	AAACGATGTCACTAGTGTGTGGGCTTATTTAG--GCTTGTGGTAAACGAACTTAAACCGGTGA	838
QY	842	GTTGACCGGCTGGGGAGTACGGTTCGCAAGATTAAACTCAAGAAATTGACGGGGACCCG	901
Db	839	GTTGACCGGCTGGGGAGTACGGTTCGCAAGATTAAACTCAATGAAATTGACGGGGGCGG	898
QY	902	CACAAGCGGTGATATATGTGGATTAATTTCCATGCAACGGGAAAACTTACTTACCCCTTG	961
Db	899	CACAAGCGGTGAGCATGTGGTTTAATTCGAAGCAACGGGAAACTTTACACAGGCTTG	958
QY	962	ACATGTAGGGAATTTTCTAGAGATAGATTAGTG--CTTCGGGAAACGCTTAAACAGGTGCTG	1020
Db	959	ACATGTAGGGAATTTTCTAGAGATAGATTAGTG--CTTCGGGAAACGCTTAAACAGGTGCTG	1018
QY	1021	CATGCGTGTCTGACGCTCGTGTCTGTAGATGTGGTTAAGTCCCGCAACGAGCGCAC	1080
Db	1019	CATGCGTGTCTGACGCTCGTGTCTGTAGATGTGGTTAAGTCCCGTAAACGAGCGCAC	1078
QY	1081	CTTGTCAATTAATTTGCCATCAATT--GTTGGGCACTTAATAGAGCTGCGGGTGAACAAC	1138
Db	1079	CTTGTCTTAATTAATTTGCCATCAATT--GTTGGGCACTTAATAGAGCTGCGGGTGAACAAC	1138
QY	1139	CGGAGGAAGGTGGGGATGATCGTCAAGTCCCTTAATGAGGTGGGCTTTCACACGT	1198
Db	1139	CGGAGGAAGGTGGGGATGATCGTCAAGTCCCTTAATGAGGTGGGCTTTCACACACGT	1198
QY	1199	AATACATATGCGCGTACAGAGGGTTGCCAACCCCGGAGGGGAGCTAATCTCAAGAACG	1258
Db	1199	GCTACATATGCGCGTACAGAGGGTTGCCAACCCCGGAGGGTGAAGCTAATCCCAACAAAC	1258
QY	1259	CGTGTATGTCGGATCGGAGTCTGCACACTGCATCTCGGAAAGTCCGGAATCGGTATATC	1318
Db	1259	GATCGTATGTCGGATCGGAGTCTGCACACTGCATCTCGGAAAGTCCGGAATCGGTATATC	1318
QY	1319	GGGATTCAGCATGTCCGGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCAAC	1378
Db	1319	GGGATTCAGCATGTCCGGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCAAC	1378
QY	1379	ATGGAGATGGGTTTCAACAGAACGAGTATGCTTAAACGTTAGAGAGGCGCTTCCACGGT	1438
Db	1379	ATGGAGATGGGTTTCAACAGAACGAGTATGCTTAAACGTTAGAGAGGCGCTTCCACGGT	1438
QY	1439	GAGATTCATGACTGGGGTG 1457	
Db	1439	GAGATTCATGACTGGGGTG 1457	
RESULT 22			
US-09-748-205-1			
Sequence 1, Application US/09748205			
Patent No. 658562			
GENERAL INFORMATION:			
APPLICANT: Canon Inc.			
TITLE OF INVENTION: Polyhydroxyalkanoate its manufacturing method, and microorganism			
FILE REFERENCE: 4351009			
CURRENT APPLICATION NUMBER: US/09/748,205			
CURRENT FILING DATE: 2000-12-27			

Query Match	83.6%	Score 1022.6	DB 3	Length 1501
Best Local Similarity	83.6%	Pred. No. 0		
Matches 1219	Conservative 0	Mismatches 234	Indels 6	Gaps 5
US-09-748-205-1				
NUMBER OF SEQ ID NOS: 1				
; SEQ ID NO 1				
; LENGTH: 1501				
; TYPE: DNA				
; ORGANISM: Pseudomonas jessenii 161 strain.				
US-09-748-205-1				
Query	3	TGAACGCTGGGCGGATGCTTTACATGCAATGCGAAGCGGACGAGCTGCTCATCT	62	
Db	1	TGAACGCTGGGCGGACGCTTACATGCAATGCGAAGCGGACGAGCTGCTCATCT	59	
Query	63	GATGCGAGTGGCGGACGGGTGAGTATGCAATCGGACGTATCCAGAAAGGGGGGTAC	122	
Db	60	GAATTC-AGCGCGGACGGGTGAGTATGCAATCGGACGTATCCAGAAAGGGGGGTAC	118	
Query	123	GCATGAAAGATGCTTAAATCCGATATCTTAAAGAGAAACAGGGGATCGAAGA	182	
Db	119	GTCTGAAAGGACGCTAATACCGATACGCTTAAAGAGAAAGGAGACCTTTGGG	178	
Query	183	CCTTGCCTTTTGGAGCGGCGGATGCTGATAGCTAGTTGGTGGGTAAAGGCTTACA	242	
Db	179	CCTTGCCTTATCATGATGAGCCATAGCTGATAGCTAGTTGGTGGGTAAAGGCTTACA	238	
Query	243	AGGCGACGATAGTATGTTGTTGTAAGAGACGACCACTGGGACTGAGACCGG	302	
Db	239	AGGCGACGATCGTAACTGTTGTTGTAAGAGACGACCACTGGGACTGAGACCGG	298	
Query	303	CCAGACTCTTACGGGAGGACGAGTGGGGAATTTTGGACAAATGGGCGCAAGCTGATCA	362	
Db	299	CCAGACTCTTACGGGAGGACGAGTGGGGAATTTTGGACAAATGGGCGCAAGCTGATCA	358	
Query	363	GCAATGCCGCGGATGAAAGAGGCTTCGGGTTGAAAGCTCTTCACTGAGAAAGA	422	
Db	359	GCAATGCCGCGGATGAAAGAGGCTTCGGGTTGAAAGCTCTTCACTGAGAAAG	418	
Query	423	AGTTACGGTAAATATATGCTGATCATGACGGTATCGACAGAAAGACCGGCTAACTA	482	
Db	419	GGCATTAACCTAATATGTTAGTTTGAAGCTTACCGACAGATTAAGACCGGCTAACTC	478	
Query	483	CGTGCACGACGCGCGGTAAATAGTAGGCTGACAGGCTTAAATCGGAATTAAGTGGGCTAA	542	
Db	479	TCGTGCACGACGCGCGGTAAATAGTAGGCTGACAGGCTTAAATCGGAATTAAGTGGGCTAA	538	
Query	543	AGGCTGCGGACGCGCTTTGTATGATCATGATGTAATCCCGGGCTTAAGCTGGGAATG	602	
Db	539	AGGCTGCGGACGCGCTTTGTATGATGTAATCCCGGGCTTAAGCTGGGAATG	598	
Query	603	CGTTTGAATCTAAGGCTAGAGTGGGACAGAGGAGTGGGAATTCATGTAGACGTG	662	
Db	599	CAATCAAACTGACAGGCTAGAGTGGGACAGAGGAGTGGGAATTCATGTAGACGTG	658	
Query	663	AAATGCGTACGATTAAGAAACATCATGAGGAGGAGCGCTCTGAGTTAACTGA	722	
Db	659	AAATGCGTACGATTAAGAAACATCATGAGGAGGAGCGCTCTGAGTTAACTGA	718	
Query	723	CGCTCATGACGAAAGCGTGGGAGGAAACAGATTAAGTACCTCTGTAGTCCAGGCCCT	782	
Db	719	CACTGAGTGGGAAAGCGTGGGAGGAAACAGATTAAGTACCTCTGTAGTCCAGGCCCT	778	
Query	783	AAACGATGCACTAGTTTGGGCTTATAG-CCTTGGTAAAGAACTTAAAGCGCTGAA	841	
Db	779	AAACGATGCACTAGTTTGGGCTTATAG-CCTTGGTAAAGAACTTAAAGCGCTGAA	838	
Query	842	GTGACCGGCTGGGAGTACGGTCCGCAAGATTAAATCTAAAGAAATGACGGGACCG	901	
Db	839	GTGACCGGCTGGGAGTACGGGCGGCAAGATTAAATCTAAATGATGACGGGACCG	898	
Query	902	CACAGCGGTGATTAATGATTAATTTGATGCAACGGGAAAACCTTAACCTACCTTG	961	

Db 899 CACAAGCGGTGAGCATGTGTTTAATTCGAAAGCAACGCGAAGAACTTACAGGCTTG 958  
Qy 962 ACATGTACGGAATTTTCTAGAGATAGATTAGT-CTTCGGGAACGCTTAACAGGTGCTG 1020  
Db 959 ACATCAATGAACCTTCCAGAGATGATGGGTGCTTCGGGAACTTGAAGACAGGTGCTG 1018  
Qy 1021 CATGCTGTCCGACCTGTGTGCGAGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1080  
Db 1019 CATGCTGTCCGACCTGTGTGCGAGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1078  
Qy 1081 CTTCGATTAATTCGATCATATT--GTTGGGCACTTTAATGAGACTGCGGTGACAAAC 1138  
Db 1079 CTTCGATTAATTCGATCATATTGATGAGCACTTAAGGAGACTGCGGTGACAAAC 1138  
Qy 1139 CGGAGAGATGGGATGACGTCAAGTCTCAATGCGCTTATGGTGGGCTTCACACGT 1198  
Db 1139 CGGAGAGATGGGATGACGTCAAGTCAATGCGCTTATGGGCTTCACACGT 1198  
Qy 1199 AATTAATGGGCGGCAAGAGGGTTGCCAACCGGAGGGAGGCTTAATCTCAGAAAGCG 1258  
Db 1199 GCTTAATGGGCGGCAAGAGGGTTGCCAACCGGAGGGAGGCTTAATCTCAGAAAGCG 1258  
Qy 1259 GCTGTAGTCCGATCGAGATCTGCAACTCGACTCGTGAAGTCGGAATCGCTAGTAATC 1318  
Db 1259 GATGCTAGTCCGATCGAGATCTGCAACTCGACTCGTGAAGTCGGAATCGCTAGTAATC 1318  
Qy 1319 GGGATCAAGCATGTGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCGCTCACAC 1378  
Db 1319 GCGAATCAAGATGTGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCGCTCACAC 1378  
Qy 1379 ATGGAGATGGGTTTCCAGCAGAGAGATAGTCTAACCGTAAGGAGGCGCTGACAGGT 1438  
Db 1379 ATGGAGATGGGTTTCCAGCAGAGAGATAGTCTAACCGTAAGGAGGCGCTGACAGGT 1438  
Qy 1439 GAGATTCATGACTGGGCTG 1457  
Db 1439 GATGATTCATGACTGGGCTG 1457

RESULT 23  
US-09-951-720-1  
; Sequence 1, Application US/09951720  
; Patent No. 6635782  
; GENERAL INFORMATION:  
; APPLICANT: Canon Kabushiki Kaisha  
; TITLE OF INVENTION: Polynucleotide and Manufacturing Method Thereof  
; FILE REFERENCE: 4477001  
; CURRENT APPLICATION NUMBER: US/09/951,720  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: JP 279900/2000  
; JP 378827/2000  
; JP 165238/2001  
; JP 165509/2001  
; JP 275063/2001  
; PRIOR FILING DATE: 2000-09-14  
; 2000-12-13  
; 2001-05-31  
; 2001-05-31  
; 2001-05-31  
; NUMBER OF SEQ ID NOS: 1  
; SEQ ID NO 1  
; LENGTH: 1501  
; TYPE: DNA  
; ORGANISM: Pseudomonas jessenii Pl61 strain.  
US-09-951-720-1

Query Match 70.2% Score 1022.6; DB 3; Length 1501;  
Best Local Similarity 83.6%; Pred No. 0;  
Matches 1219; Conservative 0; Mismatches 234; Indels 6; Gaps 5;  
Qy 3 TGAACGCTGCGCGCATGCTTTACACATGCAAGTCGAAACGCGACGACGAGTGTTCATCT 62  
Db 1 TGAACGCTGCGCGCATGCTTTACACATGCAAGTCGAAACGCGACGAGTGTTCATCT 59

Qy 63 GTGGCGAGTGGCGGACGGGATGAGTAATGATCGGAACGATCCAGAGAGAGGGGCTAAC 122  
Db 60 GAATTC-AGCGCGGACGGGATGAGTAATGCTTAGGAATCTGCTGTGATGGGGACAC 118  
Qy 123 GCATCGAAGATGTGCTAATACCGCATATACCTTAAGAGAGAAAGCAGGGATCGAAAGA 182  
Db 119 GTCTGAAAGGAGGCTAATACCGCATATACCTTAAGAGAGAAAGCAGGGATCGAAAGA 178  
Qy 183 CTTGCGCTTTTGGAGCGGCCGATGTCTGATTACTAGTTGTTGGGTTAAAGGCTTACA 242  
Db 179 CTTGCGCTTTTGGAGCGGCCGATGTCTGATTACTAGTTGTTGGGTTAAAGGCTTACA 238  
Qy 243 AGGCGAGATGATGTTGGTCTGAGAGAGACCAACGACACACTGGGACTGAGACAGCGC 302  
Db 239 AGGCGAGATGATGTTGGTCTGAGAGAGACCAACGACACACTGGGACTGAGACAGCGC 298  
Qy 303 CCAGACTCTTAACGGGAGGACAGACAGTGGGAAATTTTGAACAAATGGGCGCAACCTGATCCA 362  
Db 299 CCAGACTCTTAACGGGAGGACAGACAGTGGGAAATTTTGAACAAATGGGCGCAACCTGATCCA 358  
Qy 363 GCATTCGCGGTGATGAGAAAGCCCTTCGGGTTGTAAGCTCTTCACTGAGAGAAAG 422  
Db 359 GCATTCGCGGTGATGAGAAAGCCCTTCGGGTTGTAAGCTCTTCACTGAGAGAAAG 418  
Qy 423 AGTTACGCTAATATCTGTAATCATGACGATTCGACAGAAAGACACCGGCTTAACCTA 482  
Db 419 AGTTACGCTAATATCTGTAATCATGACGATTCGACAGAAAGACACCGGCTTAACCTA 478  
Qy 483 GGTGCCAGACGCGCGGTAATACGTAAGGTCGACAGGCTTAATCGAAATTAATCTGGGCTTA 542  
Db 479 TGTGCCAGACGCGCGGTAATACGTAAGGTCGACAGGCTTAATCGAAATTAATCTGGGCTTA 538  
Qy 543 AGGCTGGCAGGCGGCTTTGTAATCAATGTAATATCCCGGCTTAACCTGGGAAATG 602  
Db 539 AGGCTGGCAGGCGGCTTTGTAATCAATGTAATATCCCGGCTTAACCTGGGAAATG 598  
Qy 603 GCTTGAATCTAACAAGGCTAAGTGTGGCAGAGAGGAGGTGGAATTCATGTGTAGACGTG 662  
Db 599 CATTCAAAACCTGACAAAGCTAAGATGTGTAGAGGTTGGAATTTCTGTGTAGCGGTG 658  
Qy 663 AATGCTGATGATATGAGAAACATGATGATGAGGAGGAGGAGGCTCTGAGTTAACTGTA 722  
Db 659 AATGCTGATGATATGAGAAACATGATGATGATGAGGAGGAGGAGGCTCTGAGTTAACTGTA 718  
Qy 723 GCTTCATGCAAGAAAGCTGGGAGCAAAACGATTAAGTACCTGATGTCACGCGCT 782  
Db 719 CACTGAGGTGCAAAAGCTGGGAGCAAAACGATTAAGTACCTGATGTCACGCGCT 778  
Qy 783 AAACGATGCAACTAGTTGTGGGCTTATTAG-GCTTGTGTAAGAAAGCTAACCGCTGAA 841  
Db 779 AAACGATGCAACTAGTTGTGGGCTTATTAGGCTTATTAGGCGGACGTAACGATTAA 838  
Qy 842 GTTACCGCTGGGAGATGCGTGCAGAGATTAAATCAAGAAATTAAGCGGAGACCG 901  
Db 839 GTTACCGCTGGGAGATGCGTGCAGAGATTAAATCAAGAAATTAAGCGGAGACCG 898  
Qy 902 CACAAGCGGTGAGATTATGATTAATTCGATGCAACGCGAAACCTTACTACCTCTG 961  
Db 899 CACAAGCGGTGAGATGATGATTAATTCGATGCAACGCGAAACCTTACTACGAGCTTG 958  
Qy 962 ACATGTACGGAATTTTCTAGAGATAGATTAGT-CTTCGGGAACGCTTAACAGGTGCTG 1020  
Db 959 ACATCAATGAACCTTCCAGAGATGATGGGTGCTTCGGGAACTTGAAGACAGGTGCTG 1018  
Qy 1021 CATGCTGTCCGACCTGTGTGCGAGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1080  
Db 1019 CATGCTGTCCGACCTGTGTGCGAGATGTTGGGTTAAGTCCGCAACGAGCGCAAC 1078  
Qy 1081 CTTCGATTAATTCGATCATATT--GTTGGGCACTTTAATGAGACTGCGGTGACAAAC 1138  
Db 1079 CTTCGATTAATTCGATCATATTGATGAGCACTTAAGGAGACTGCGGTGACAAAC 1138

QY	1139	CGAGGAAGTGGGGATACGTCAAGTCTCATGGCCTTATGGGTAGGGCTTACACGT	1198
Db	1139	CGAGGAAGTGGGGATACGTCAAGTCTCATGGCCTTATGGGTAGGGCTTACACGT	1198
QY	1199	AATCAATAGGGCGCGTACAGAGGGTGGCCACCCGGAGGGGAGCTAATCCAGAAACG	1258
Db	1199	GCTACAAATGGTCGGTACAGAGGGTGGCCACCCGGAGGGTGAAGTAAATCCACAAAC	1258
QY	1259	CGTGTAGTCCGGATCGGAGTCTGCACATCGATCCGTGAAGTGGGAATGCTAGTATC	1318
Db	1259	GATGTAGTCCGGATCGGAGTCTGCACATCGATCCGTGAAGTGGGAATGCTAGTATC	1318
QY	1319	GCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCCGTCACAC	1378
Db	1319	GCGAATCAGAAATGTGCGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCCGTCACAC	1378
QY	1379	ATGGAGTGGGTTTACACGAAGCAGTAACTAACCGTAAAGAGGGGCGTTGCACCGGT	1438
Db	1379	ATGGAGTGGGTTTACACGAAGTAACTAACTAACCGTAAAGAGGCGTTTACACCGGT	1438
QY	1439	GAGATTCAATGACTGGGGTG	1457
Db	1439	GTGATTCAATGACTGGGGTG	1457

```

RESULT 24
US-10-411-319-1
; Sequence 1, Application US/10411319
; Patent No. 6649381
; GENERAL INFORMATION:
; APPLICANT: Canon Inc.
; TITLE OF INVENTION: Polyydroxynate, Method For Production Thereof And Microorganisms
; TITLE OF INVENTION: In The Same
; FILE REFERENCE: 03500.015001.1
; CURRENT APPLICATION NUMBER: US/10/411.319
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: US 09/748,205
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jesseni 161 strain
US-10-411-319-1

```

Query Match	70.2%	Score 1022.6;	DB 3;	Length 1501;
Best Local Similarity	83.6%	Pred. No. 0;		
Matches 1219; Conservative	0;	Mismatches 234;	Indels 6;	Gaps 5

Qy	3	TCGACGCTGGCGGCGCATGCTTTTACATCGAATCGAACGGGACGAGATGGCTTCATCT	62
Db	1	TCGAACGCTGGCGGCGCGCTTAAACAATGCAATCGACGGATG -ACGGAGCTTCTCTCT	59
Qy	63	GGTGGCGAGTGGCGGACGGGCTGAGTAATCATCGGAACGTATCCAGAAAGGGGGGTAAAC	122
Db	60	GAAATTC -ACCGGGGACGGGGTAGTAATCTCTAGGAATTTGCTGGTAGTGGGGGACAAC	118
Qy	123	GCATCGAAGATGTGCTTAATACCGCATATACTCTAAGAGGAAGACAGGGGATCGAAGA	182
Db	119	GTCTCGAAGGAGACGCTTAATACCGCATATAGCTCTACGGGAGAAAGAGGGGACCTTTGGGG	178
Qy	183	CGTTGGCTTTTGGAGCGGCGGATGTCGTATTAGCTAGTGGTGGGGTAAAGGCTTACCA	242
Db	179	CGTTGGGCTATCGAATGAGGCTTAGGTGCGGATTAGTAGTTGGTGAAGTAAATGGCTCACCA	238
Qy	243	AGGCGACGATCAGTAGTTGGTCTGAGAGGACGACAGCCACACTGGGACTTGAGACACGGC	302
Db	239	AGGCGACGATCCGTAACGTGGTCTGAGAGGATGATCAGTCAACTGGAACTGAGACACGGT	298
Qy	303	CCAGACTCTCTACGGGAGGCGACAGTGGGGAATTTTGGACAAATGGGGCGAAAGCTGATCCA	362
Db	299	CCAGACTCTCTACCGGAGGCGACAGTGGGGGAATTTGGAACAATGGGGGAAGGCTTGAATCCA	358

OY	363	GCAATGCCCGGTGAGTGAAGAAGCCCTTCGGGTTGTAAACTCTTTCAGTCGAGAAATA	422
Db	359	GCCATGCCCGGTGTGTGAMAAAGCTCTTCGATTGTAAACAATTAAAGTTGGAGAGAC	418
OY	423	AGGTTACGGTAAATATATCGTACCTCATGACGGTATCGACAGAAAGACACCGGCTAACTA	482
Db	419	GGCATTAACTTAATACGTTAGTGTGTTGACGTTACCGACGAATTAACACACCGGCTAACTC	478
OY	483	CGTCCACGACACCGCGGTAAATCGTAGGGTCAAGCGTTAATCGGAATTACTGSGGCTAA	542
Db	479	TGTCCACGACACCGCGGTAATACAGAGGATCAAGCGTTAATCGGAATTACTGSGGCTAA	538
OY	543	AGGTCGGCAGCGCGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACCTGGGAATTG	602
Db	539	AGCCCGGTAGTGTGTTTAAGTTGAATGTGAACCCCGGGCTCAACTGSGGAACCTG	598
OY	603	CGTTTGAAACTCAAGGCTAGAGTGTGGCAGAGGGAAGTGGAAATTCATGTGTAGCAGTG	662
Db	599	CATTCAAACTGACACAGCTAGATGATGTAGAGGATGGTGGAAATTTCTGTGTAGCGGTG	658
OY	663	AAATGCGTAGAGATATGGAGAAACATCGATGGCGAAGCAGCCTCTGGGTTAACTGTGA	722
Db	659	AAATGCGTAGATATAGGAAGAAACACAGATGGCGAAGCAGCCTGTGATATCTGA	718
OY	723	CGCTCATGACGAAACCGTGGGAGCAAAACGAAATTAGATACCTCGTGTATGTCACGCCCT	782
Db	719	CACGTAGGTGCGAAAGCGTGGGAGCAAAACGAAATTAGATACCTCGTGTATGTCACGCCGT	778
OY	783	AAACGATGTCACTAGTTGTTGGGCTTATTAG-GCTTGTAAGAAAGCTAACGGGTGA	841
Db	779	AAACGATGTCACTAGCGGTGGAGCCCTTAGCTTTAGTGGCGACACTAACGCATTAA	838
OY	842	GTTGACCGCTGGGGAGTACGCTGCGCAAGTTTAAACTCAAGGAATTGACGGGAGCCG	901
Db	839	GTTGACCGCTGGGGAGTACCGCGCAAGTTTAAACTCAATGATTTGACGGGGCCG	898
OY	902	CACAAGCGGTGATTTATGTGATTAAITGATGCAACGCGAAAACTTACTACCTTCG	961
Db	899	CACAAGCGGTGAGCATGTGTTTAITCGAAGCAACCGAAGAACTTACCAAGGCTTG	958
OY	962	ACATGTACGGAATTTCTAGAGATTGATTAGTG-CTTGGGGAAGCTTAACAGGTGCTG	1020
Db	959	ACATCAATGAACTTTCAGAGATGAGTGGTGCTTGGGAAACATTAGACAGGTGCTG	1011
OY	1021	CATGCTGTCGACGCTGTCGTGTCGAGATGTTGGTAAAGTCCCGCAACGACGCAAC	1080
Db	1019	CATGCTGTCGACGCTGTCGTGAATGTGGTTAAAGTCCCGTAACGAGGCAAC	1076
OY	1081	CTTGTCAATTAATTTGCCATCAATT--GGTTGGGACTTTAATGACATCGCCGCTGCAAAAC	1133
Db	1079	CTTGTCTTAATTTAGTTACACAGACGTAATGGTGGGACTTAAGGAGACTCGCGGTGCAAAAC	1130
OY	1139	CGAGGAAGGTGGGATGACGTCAAGTCTCATGGCCCTTAATGGTAAAGGCTTCAACAGT	1199
Db	1139	CGAGGAAGGTGGGATGACGTCAAGTCAATGACCTTACCGGCTTGACGCTTGACACAGT	1199
OY	1199	AATTAACAATGGCGGTACAGAGGGTGGCAACCGCGAGGGGAGCTTAATCTCAGAAAGCG	1258
Db	1199	GCTACAATGTGTGTAACAGAGGTTTGCACAGCCGGAAGTGAAGCTTAATCCACAAAAAC	1255
OY	1259	CGTGTAGTCGCGATTCGAGTCTGCACTCGACTCGTGAAGTGGAAATCGCTAGTAATTC	1311
Db	1259	GATTCGATGTCGGAATCGCAGTCTGCACATCGACTCGTGAAGTGGAAATCGCTAGTAATTC	1311
OY	1319	GGCGATCAGCATGTGCGCGGTGAATACGTTCCCGGGTCTTTGTACACACCGCCCGTCAACC	1376
Db	1319	GGCAATCAAAATGTGCGCGGTGAATACGTTCCCGGGCTTTGTACACACCGCCCGTCAACC	1376
OY	1379	ATGGAGATGGGTTTACACGAAGACAGATGTAAACGTAAGGAAGGCGCTTGACACGGT	1433
Db	1379	ATGGAGATGGGTTTACCAAGAAATGACTAGTAACTTTCCGAGAGACGGTTTACACGGT	1433







```
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii P161 ; Bp-7376
; FEATURE:
; FEATURE: cDNA to 16S rRNA
US-10-266-787-5

Query Match      70.2%; Score 1022.6; DB 3; Length 1501;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 1219; Conservative 0; Mismatches 234; Indels 6; Gaps 5;

QY      3 TGAACGCTGGGCGGAGCTTTACATGCAAGTCGAAACGGACGACGAGTGTTCATCT 62
DB      1 TGAACGCTGGGCGGAGCTTTACATGCAAGTCGAAAGTCGAGCGATG-ACGGAGCTTGTCTCT 59
QY      63 GGTGCGAGTGGCGGACGGGTGAGTAATGCATCGAAACGTATCCAGAAAGGGGGTAAAC 122
DB      60 GAATTC-AGCGGCGGACGGGTGAGTAATGCTAGGAATCTGCTGGTAGTGGGGGACAAAC 118
QY      123 GCATGAAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGACAGGGGATCGAAAGA 182
DB      119 GTCTGAAAGGAGCCTTAATACCGCATATGCTCTACGGGAGAAAGAGGGGACCTTCGGG 178
QY      183 CTTGCGCTTTTGAAGCGGCGCATGCTGATTAGCTAGTTGGTGGGTTAAAGGCTTACCA 242
DB      179 CTTGCGCTATCATGATGAGCTTAAGTCTGATTAAGTCTAGTTGGTGGGTTAAAGGCTTACCA 238
QY      243 AGGCGACATCATAGTGTGCTGAGAGACGACCAACGACCACTGGGACTGAGACACGGC 302
DB      239 AGGCGACATCATGATGCTGATGCTGAGAGATGATGATGATGATGATGATGATGATGATGAT 298
QY      303 CCAGACTCTTACGAGGAGGACGAGATGTTGGAAATTTTGAACAATGGGGGCAAGCTTATCCA 362
DB      299 CCAGACTCTTACGAGGAGGACGAGATGTTGGAAATTTTGAACAATGGGGGCAAGCTTATCCA 358
QY      363 GGAATGCGCGGTGAGAGAGAGGCGCTTGGGGTGTAAAGCTCTTCAGTCGAGAGAAA 422
DB      359 GCAATGCGCGGTGAGAGAGAGGCTTGGGGTGTAAAGCTCTTCAGTCGAGAGAAAG 418
QY      423 AGGTTACGTTAATATCGTGACTCATGACGATTCGACAGAGAACACCGGCTTACTA 482
DB      419 GGCATTAACTTAATATAGTATAGTGTGTAAGTTAACGACAGAAATAGACACCGGCTTAATC 478
QY      483 CGTGCCAGACGCGCGGTAAATACGTAAGGCTGCAAGCTTAATTCGAATTAATCTGGCGCTAA 542
DB      479 TGTGCCAGACGCGCGGTAAATACGTAAGGCTGCAAGCTTAATTCGAATTAATCTGGCGCTAA 538
QY      543 AAGGTGCGGAGGCGGCTTTGTAGTACGATGTAATCCCGGGCTTAATCTGGGAATGG 602
DB      539 AAGGTGCGGAGGCGGCTTTGTAGTACGATGTAATCCCGGGCTTAATCTGGGAATGG 598
QY      603 CGTTTGAACCTACAGGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGAGAGTG 662
DB      599 CATTTAAACTGACAGGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTAGAGAGTG 658
QY      663 AAATGCGTATGATATGAGAAACATGATGCGAAAGGACGCTCTGGGTTAACTGA 722
DB      659 AAATGCGTATGATATGAGAAACATGATGCGAAAGGACGCTCTGGGTTAACTGA 718
QY      723 CGGTATGACGAAACGCTGGGAGGAAACAGGATTAAGATCCCTGGTATCCACGCCCT 782
DB      719 CACTGAGGTGCGAAAGGTGGGAGCAACAGGATTAAGATCCCTGGTATCCACGCCCT 778
QY      783 AAACGATGCTCACTAGTTGTGGGCTTATTAAG-GCTTGGTAAAGAACTAAAGCGCGTAA 841
DB      779 AAACGATGCTCACTAGCCGTTGGGAGCGCTTGAAGCTTATTAAGCGGACGCTAAAGCAATTA 838
QY      842 GTTACCGGCTGGGAGTACGCTGCGCAAGTAAATCAAGGAATTTGACGGGACCGG 901
```

```
DB      839 GTTACCGGCTGGGAGTACGCGCGCAAGGTTAAACTCAATGATTAAGACGGGGCCCG 898
QY      902 CACAAGCGGTGATTAATGTGATTAATTCATGCAAGCGGCAAAACCTTACTTACCTCTG 961
DB      899 CACAAGCGGTGATTAATGTGATTAATTCATGCAAGCGGCAAAACCTTACTTACCTCTG 958
QY      962 ACATGTAGCGAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTTAACACAGGTGCTG 1020
DB      959 ACATTCATTAATCTTCCAGAGATGATGAGTGTGCTTCGGGAAACATTAAGACAGGTGCTG 1018
QY      1021 CATGCTGTCTGCAAGCTGCTGCTGAGATGTTGGGTTAACTCCCGCAACAGCGCAAC 1080
DB      1019 CATGCTGTCTGCAAGCTGCTGCTGAGATGTTGGGTTAACTCCCGCAACAGCGCAAC 1078
QY      1081 CTTGTCTTAATTTGCGATCATTTT--GTTGGGCACTTAAATGACATGCGCGGTGCAAAAC 1138
DB      1079 CTTGTCTTAATTTGCGATCATTTT--GTTGGGCACTTAAATGACATGCGCGGTGCAAAAC 1138
QY      1139 CGGAGGAAAGGTGGGATGACGTCAGATCTCATGAGCCCTTAATGGGTAGGGCTTCACAGT 1198
DB      1139 CGGAGGAAAGGTGGGATGACGTCAGATCTCATGAGCCCTTAATGGGTAGGGCTTCACAGT 1198
QY      1199 AATACATGCGCGCTACAGAGGTTGCCAACCCGCGAGGAGGAGCTTAATTCAGAAAGCG 1258
DB      1199 GCTACATGATGCTGATGACAGAGGTTGCCAACCCGCGAGGAGGAGCTTAATTCAGAAAGCG 1258
QY      1259 CGTGTAGTCCGAGATCGAGATCTGCACTCCGTTGAAAGTGGGAATTCGCTAGTAATC 1318
DB      1259 GATCTAGTCCGAGATCGAGATCTGCACTCCGTTGAAAGTGGGAATTCGCTAGTAATC 1318
QY      1319 GCGGATGACGATGTCGCGGTGATGCTTCCGGGCTTTGTACACACCGCCGTCACACC 1378
DB      1319 GCGGATGACGATGTCGCGGTGATGCTTCCGGGCTTTGTACACACCGCCGTCACACC 1378
QY      1379 ATGGAGTGGGTTTACAGAGAGCAGTATGCTTAAACGTAAGAGGCGGCTTCCACGCT 1438
DB      1379 ATGGAGTGGGTTTACAGAGAGCAGTATGCTTAAACGTAAGAGGCGGCTTCCACGCT 1438
QY      1439 GAGATTCACTGCTGGGGTG 1457
DB      1439 GTGATTCATGACTGGGGTG 1457

RESULT 27
US-09-791-610-1
; Sequence 1, Application US/09791610
; Patent No. 6861550
; GENERAL INFORMATION:
; APPLICANT: Canon Inc.
; TITLE OF INVENTION: Polyhydroxyalkanoate containing 3-hydroxybenzoylalkanoic acid as
; TITLE OF INVENTION: monomer unit, and method for producing the same.
; FILE REFERENCE: 4396021
; CURRENT APPLICATION NUMBER: US/09/791,610
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii 161 strain.
US-09-791-610-1

Query Match      70.2%; Score 1022.6; DB 3; Length 1501;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 1219; Conservative 0; Mismatches 234; Indels 6; Gaps 5;

QY      3 TGAACGCTGGGCGGAGCTTTACATGCAAGTCGAAACGGACGACGAGTGTTCATCT 62
DB      1 TGAACGCTGGGCGGAGCTTTACATGCAAGTCGAAAGTCGAGCGATG-ACGGAGCTTGTCTCT 59
QY      63 GGTGCGAGTGGCGGACGGGTGAGTAATGCATCGAAACGTATCCAGAAAGGGGGTAAAC 122
DB      60 GAATTC-AGCGGCGGACGGGTGAGTAATGCTAGGAATCTGCTGGTAGTGGGGGACAAAC 118
```

```
QY 123 GCATCGAAGAGATGTCCTAATACCGCATATACCTAAGAGAGAAAGCAGGGGATCGAAGA 182
Db 119 GCTCGAAGAGAGACCTAATATACCGCATACGTCCTACGGAGAGAAAGGAGACCTTCGAG 178
QY 183 CTTTCGCGTTTGGAGCGCGCATGTCGATTTAGCTAGTTGGTGGGTTAAAGGCTTACCA 242
Db 179 CTTTCGCGTTATGAGATGAGCCCTAGGCTTGGATTAGCTAGTTGGTGGGTTAAAGGCTTACCA 238
QY 243 AGGCGACGATCATGATGTTGGTCTGAGAGACGACCAAGCCTGAGCTGAGACACAGGC 302
Db 229 AGGCGACGATCCGTTAATCGTCTGAGAGATGATCATGACACTGGAACCTGAGACACGCT 298
QY 303 CCAAGCTCTTACCGGAGGACGAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATCCA 362
Db 299 CCAAGCTCTTACCGGAGGACGAGTGGGGAATTTTGGACAATGGGCGCAAGCTTATCCA 358
QY 363 GCAATGCGCGTGAATGAGAGAGCCTTCGAGTTTAAAGCTCTTTCAGTCGAGAGAAA 422
Db 359 GCAATGCGCGTGTGTGAAGAGAGCTTCGAGTTTAAAGCTTTCAGTTGGAGAGAG 418
QY 423 AGGTTACGGTAATATATCTGTGATCATGACGGTATCGACAGAGAGACCGGGCTAATCA 482
Db 419 GCGATTAACTAATATCGTTAGTGTTTGACGTTACCGACAGAAATAGGACCGGCTAATCTC 478
QY 483 CGTGCACGACCGCGGGTAAATACGTAAGGTGCAAGGTTAAATCGGAATTTAATCGGGCTTAA 542
Db 479 TGTGCACGACCGCGGGTAAATACGTAAGGTGCAAGGTTAAATCGGAATTTAATCGGGCTTAA 538
QY 543 AGGCGCGCAGCGCGCTTTGTAGTCAATGATGAATCCCGGGGCTTAACTGAGAAATTCG 602
Db 539 AGCGCGCGTATGAGTGTGTTTGAATGTTGAATGGAAGAGCCCGGGCTCAACCTGAGAACTG 598
QY 603 CGTTTGAATCTAACAAGCTAGAGTGTGCAAGGAGGTGGAATTCATGTGTAGCAAGT 662
Db 599 CATTCAAAATCGAACAGCTAGAGTGTGAGAGGCTGTGGAAATTCCTGTGTAGCGGTG 658
QY 663 AAATGCGTAAGATATGGAAGAACATCGATGCGGAAGGACGCGCTGCTGTTAAACATCGA 722
Db 659 AAATGCGTAAGATATGGAAGAACACAGTGTGCGAAGGACGCGCTGCTGATGATCTGA 718
QY 723 CGCTCATCAGCAAGAAAGCGTGGGAGCAAAACAGAGTTAGTAACTCTGTAGTCAAGCCCT 782
Db 719 CACTGAGGTGCGAAGAAAGCGTGGGAGCAAAACAGAGTTAGTAACTCTGTAGTCAAGCCCT 778
QY 783 AAACGATGTCAACTAGTGTGTGGGCTTAATAG-GCTTGTGTAAGCAAGCTAACCGGTGA 841
Db 779 AAACGATGTCAACTAGCGGTGTGGAGCCCTTGAAGCTTAATGTGGCCAGCTAACCGCATTA 838
QY 842 GTTGAACCGGCTGGGAGATACGCTCGCAAGATTTAAACTCAAGGAATTGACGGGGAGCCG 901
Db 839 GTTGAACCGGCTGGGAGATACGCGCGCAAGGTTAAACTCAAGGAATTGACGGGGAGCCG 898
QY 902 CACAAGCGGTGGATATATGATTAATTCGATGCAACGCGAATACTTAACCTTAACCTTTCG 961
Db 899 CACAAGCGGTGGAGATGTGTGTTAATTCGAGAACAGCGAAGAACTTAACCAAGGCTTCG 958
QY 962 ACATGTAGCGAATTTTCTAGAGATGATTAAGTG-CTTGGGAAACGCTTAACACAGGTCTG 1020
Db 959 ACATGTAGCGAATTTTCTAGAGATGATTAAGTGCTTGGGTTAAAGTCCGTTAACGAGGTCTG 1018
QY 1021 CATGCGTGTGCTACACTCGTGTGCTGAGATGTTGGGTTAAAGTCCGCAACGAGCGCAAC 1080
Db 1019 CATGCGTGTGCTACACTCGTGTGCTGAGATGTTGGGTTAAAGTCCGTTAACGAGCGCAAC 1078
QY 1081 CTTGTCAATTAATTTGCAATCATTT--GGTTGGGCACTTTAATGAGACTGCGGCTGCAAA 1138
Db 1079 CTTGTCTTATAGTTACAGACGCTAATGTGTGGCACTTAAGAGAGACTGCGGCTGCAAA 1138
QY 1139 CGAGAGAGGTGGGATGACGTCAAGTCTCATGGCCCTTAATGGGATGGGCTTCAACAGCT 1198
Db 1139 CGAGAGAGGTGGGATGACGTCAAGTCTCATGGCCCTTAATGGGCTTCAACAGAGCT 1198
QY 1199 AATACAAATGGCGGCTACAGAGGTTGCCAACCCGAGAGGGGAGCTAATCTCAGAAAGCG 1258
```

```
Db 1199 GCTACAAATGTCGCGATACAGAGGTTGCCAAGCCGCGAGGTGAGACTAATCCACAAAACC 1258
QY 1259 GGTGTAGTCCGGAATCGGAGTCTTCAACTGCACTTCCTGGAAGTGGGAATCGCTGTAATC 1318
Db 1259 GATGTAGTCCGGAATCGGAGTCTTCAACTGCACTTCCTGGAAGTGGGAATCGCTGTAATC 1318
QY 1319 GCGGATCAGCATGTCGCGGTAATACGTTCCCGGCTCTTGTACACACCGCCGTCACACC 1378
Db 1319 GCGAATCAGAAATGTGCGGGTAATATGTTCCCGGCTCTTGTACACACCGCCGTCACACC 1378
QY 1379 ATGGAGTGGGTTTACCCAGAGAGGATGATCTAATCGTAAAGAGGCGCTTTCACGCT 1438
Db 1379 ATGGAGTGGGTTTACCCAGAGAGGATGATCTAATCGTAAAGAGGCGCTTTCACGCT 1438
QY 1439 GAGATTCAATGACTGGGGTG 1457
Db 1439 GTGATTCAATGACTGGGGTG 1457
```

```
RESULT 28
US-10-252-518-5
; Sequence 5, Application US/10252518
; Patent No. 6875596
; GENERAL INFORMATION:
; APPLICANT: Yano, Tetsuya
; APPLICANT: Imamura, Takeshi
; APPLICANT: Suda, Sakae
; APPLICANT: Homma, Tetsuomu
; TITLE OF INVENTION: Polyoxyalkanoate Synthase and Gene Encoding the Same Enzyme
; FILE REFERENCE: 03500.015225.2
; CURRENT APPLICATION NUMBER: US/10/252,518
; CURRENT FILING DATE: 2002-09-24
; PRIOR FILING DATE: 2000-095004
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii Pl61 ; BP-7376
; FEATURE:
; FEATURE: cDNA to 16S rRNA
US-10-252-518-5
```

```
Query Match 70.2%; Score 1022.6; DB 3; Length 1501;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 1219; Conservative 0; Mismatches 234; Indels 6; Gaps 5;
```

```
QY 3 TGAACGCTGGGCGGATGCTTTAACAATGCAATGGAACCGGACGAGCTGTCATCT 62
Db 1 TGAACGCTGGGCGGAGCTTAACAATGCAATGGAACCGGACGAGCTGTCATCT 59
QY 63 GGTGGCGAGTGGCGGAGCGGTGAGTAATGCAATCGGAACTGATCCAGAGAGGGGTTAAC 122
Db 60 GAATTC-AGCGCGGAGCGGTGAGTAATGCTTAAGAAATCTGCTGTGTGTGGGGAGAAC 118
QY 123 GCATCGAAGAGTGTGCTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGGATGAAAGA 182
Db 119 GCTTCGAAGAGGAGCTAATACCGCATTAATCTTAAGAGAGAAAGCAGGGGATGAAAGA 178
QY 183 CTTGCGCGTTTGGAGCGGCGGATGCTGATTAAGTGTGGTGGGTTAAAGGCTTACCA 242
Db 179 CTTGCGCTATCAATGAGCTTAAAGCTGATTAAGTGTGGTGGGTTAAAGGCTTACCA 238
QY 243 AGGCGACGATCAGTATGTTGTCTGAGAGAGACGACCACTGAGCTGAGACACGCGC 302
Db 239 AGGCGACGATCCGTAATCTGCTGAGAGAGATGATCACTGAGAACTGAGAACAGCT 298
QY 303 CCAAGCTCTTACCGGAGGACAGTGGGAAATTTTGGCAAATGGGCGCAAGCTGATCCA 362
Db 299 CCAAGCTCTTACCGGAGGACAGTGGGAAATTTTGGCAAATGGGCGCAAGCTGATCCA 358
```

```

Qy 363 GCATGCCGCGTGAAGAAAGCCCTTCGGGTTGTAAGCTCTTCACTCGAAGAA 422
Db 359 GCCATGCCGCGTGTGAAGAGGTCTTCGATTGAAAGCACTTAAATTGGAGAG 418
Qy 423 AGGTTACGGTAATATCGTCACTGACGGTATCGACAGAAAGACCGGCTAACTA 482
Db 419 GGCATTAACTTAATACGTTAGTGTGTTGACGTTACCGACAGAAATAGCACCGGCTACTC 478
Qy 483 CGTGCAGACGCGCGGTAAATACGTAAGGTGCAAGCGTTAATCGAAATTAATCTGGCGTAA 542
Db 479 TGTGCTAGAGCGCGGTAAATACAGAGGTGCAAGGTTAATCGGAATTAATCTGGCGTAA 538
Qy 543 AGGGTCCGACGCGCGCTTGTAAAGTCAGATGTGAATCCCGGGCTTAATCTGGGAATG 602
Db 539 AGCGCGGTAGTGTGTGTTGTTAAAGTTGATGTAAGCCCGGCTCAACCTGGGAACCTG 598
Qy 603 CGTTGAAACTACAGGCTAGAGTGTGAGAGGGAGGAGGTAATTCATGTAGAGAGT 662
Db 599 CATTGAAACTACAGGCTAGAGTGTGAGAGGGAGGAGGTAATTCATGTAGAGAGT 658
Qy 663 AAATGCGTAGATATGAAGAAACATCGATGCGAAGGCGAGCGCTCTGGGTTAACACTGA 722
Db 659 AAATGCGTAGATATGAAGAAACATCGATGCGAAGGCGAGCGCTCTGGGTTAACACTGA 718
Qy 723 CGCTCATGACAGAAAGCGTGGGAGCAACAGGATTAGATACCTGTGTAGTCCAGCCCT 782
Db 719 CACTGAGGTGCGAAAGCGTGGGAGCAACAGGATTAGATACCTGTGTAGTCCAGCCCT 778
Qy 783 AAACGATGTCAACTAGTGTGGGCTTATAG-GTTGTGTAACGAAGTAAAGCGCTGA 841
Db 779 AAACGATGTCAACTAGTGTGGGCTTATAG-GTTGTGTAACGAAGTAAAGCGCTGA 838
Qy 842 GTTGAACCGCTGGGAGTACGGTCCGAAAGTTAACTCAAAGAAATTAACGGGAGCCG 901
Db 839 GTTGAACCGCTGGGAGTACGGTCCGAAAGTTAACTCAAAGAAATTAACGGGAGCCG 898
Qy 902 CACAAGCGGTGATTAATGTGATTAATTCGATGCAACGGGAAACCTTAACCTTCTG 961
Db 899 CACAAGCGGTGATTAATGTGATTAATTCGATGCAACGGGAAACCTTAACCTTCTG 958
Qy 962 ACATGTAGCGAATTTCTAGAGTAAATTAATAGT-CTTCGGGAACGCTAACAGGTCG 1020
Db 959 ACATGTAGCGAATTTCTAGAGTAAATTAATAGT-CTTCGGGAACGCTAACAGGTCG 1018
Qy 1021 CATTGCTGTGACGCTCGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACC 1080
Db 1019 CATTGCTGTGACGCTCGTGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACC 1078
Qy 1081 CTTGTCTAATTAATTCATCATTT--GTTGGGCACTTAAATGAGCTGCGGTGACAAAC 1138
Db 1079 CTTGTCTAATTAATTCATCATTCATGATGTTGGGCACTTAAATGAGCTGCGGTGACAAAC 1138
Qy 1139 CCGAGGAAGGTGGGATGACGTCATGCTCTCATGCGCTTATAGGCTTCAACAGT 1198
Db 1139 CCGAGGAAGGTGGGATGACGTCATGCTCTCATGCGCTTATAGGCTTCAACAGT 1198
Qy 1199 AATACAAATGCGGTGACAGAGGTGCAACCGCGAGGGGAGCTAATCTCAGAAAGCG 1258
Db 1199 GCTACAAATGCGGTGACAGAGGTGCAACCGCGAGGGGAGCTAATCTCAGAAAGCG 1258
Qy 1259 CGTCTAGATCCGATCGAGTCTGCAACTGACTCCGTGAAAGTCCGAATGCTAGTAATC 1318
Db 1259 GATCTAGATCCGATCGAGTCTGCAACTGACTCCGTGAAAGTCCGAATGCTAGTAATC 1318
Qy 1319 GCGGATCAGCATGTCCGGGTGAATACGTTCCCGGGTTTGTACACACCGCCCTGACACC 1378
Db 1319 GCGGATCAGCATGTCCGGGTGAATACGTTCCCGGGTTTGTACACACCGCCCTGACACC 1378
Qy 1379 AATGAGTGGGTTTCAACAGAGAGGTAGCTTAACCGTAAAGGAGGGCGCTGCAAGT 1438
Db 1379 AATGAGTGGGTTTCAACAGAGAGGTAGCTTAACCGTAAAGGAGGGCGCTGCAAGT 1438
Qy 1439 GAGATTCATGACTGGGCTG 1457

```

```

Db 1439 GTGATTCATGACTGGGCTG 1457

RESULT 29
US-08-632-470-53
; Sequence 53, Application US/08632470
; Patent No. 5976791
; GENERAL INFORMATION:
; APPLICANT: MABILAT, CLAUDE
; APPLICANT: RAOUIT, DIDIER
; TITLE OF INVENTION: NUCLEOTIDE FRAGMENTS CAPABLE OF
; HYBRIDIZING SPECIFICALLY TO RICKETTSIA RNA OR RNA AND
; TITLE OF INVENTION: THEIR USE AS PROBES OR PRIMERS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIVIER & BERRIDGE
; STREET: P. O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/632,470
; FILING DATE: 08-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPA 38238
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)-836-6400
; TELEFAX: (703)-836-2787
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1484 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-632-470-53

Query Match 70.1%; Score 1020.8; DB 2; Length 1484;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 243; Indels 7; Gaps 4;

Qy 1 ATTGAAGCTGCGCGGATCTTTTACATGCAAGTGAACGCGACAC--GGATGCTTGC 58
Db 28 ATTGAAGCTGCGCGGATCTTTTACATGCAAGTGAACGCGACGCGAGGAGNCTTNC 87
Qy 59 ATCTGTGAC--GAGTGCAGACGCGGTGATGATGCAACGATTCACAAAGAGGGG 116
Db 88 TCCTGCGCGCAGAGAGTGGCGGACGCGGTGATGATGCGATGCAATCTTGTAGTGGG 147
Qy 117 GGTAAAGCATGAAATATGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGAGTC 176
Db 148 GATTAACCTGGGAAACTCGGGCTAATACGTAATATCTTTGGAACAAAGCGGGGATC 207
Qy 177 GAAAGACTTTCGCTTTTGGAGCGCGCATGTCGTGATTAGCTAGTGTGGGGTAAAGC 236
Db 208 TTGGAACCTTCGTCTTAAGATGAGCTTAAGTGAATGCTTGTGTGGGGTAAATGCG 267
Qy 237 CTACCAAGCGCAGATCAGTATGTTGTTGTAAGAGAGCAACGCCACACTGGGACTGAGA 296
Db 268 CTACCAAGCGCAGATCAGTATGTTGTTGTAAGAGAGCAACGCCACACTGGGACTGAGA 327
Qy 297 CACGCGCCAGATCTCTACCGGAGGACGAGTGGGGAATTTTGGCAATGGGCGCAAGCT 356

```

Db 328 CNGGCGCCAGACTCTCTACGGGAGGAGCAGCATGGGGATATTGGACATATGGGGGAAACCT 387  
Qy 357 GATCCAGCAATGCCCGCTGAGTGAAGAGGCTTGGGTTGTAAAGCTTTCAATCGAG 416  
Db 388 GATCCAGCAATGCCCGCTGAGTGAAGAGGCTTGGGTTGTAAAGCTTTCAATCGAG 447  
Qy 417 AAGAAAGGTTACGGTAAATATCGTGAATCATGAGTATGACAGAAAGACACCGG 476  
Db 448 AAGAAATTTCTCAGGGTAAATATCTTGGGCGCTTGAAGCTTACCAAGAAAGACCTGG 507  
Qy 477 TAACTACGTGCAGAGCGCGGTAAATATGTAAGGTGCAAGGCTTAATCGAATTACTGG 536  
Db 508 TAACTCTGTGCAGAGCGCGGTAAATATGTAAGGTGCAAGGCTTAATCGAATTACTGG 567  
Qy 537 GCGTAAAGGTGCGCAGCGGCTTTGTAAGTGAATGTAATCCCCGGGCTTAACCTGG 596  
Db 568 GCGTAAAGCGCGCTGAGGTGATATTGAAGTGTGAAGAGCCCTGGGCTNNACCTGG 627  
Qy 597 GAATTGGCTTTGAATCAAGGCTAAGTGTGAGAGGAGGAGGTAATTCATGTGTA 656  
Db 628 GAATTGCAACCGATATCTGGGTATCTTGAAGTATGTAAGGAAATGGAATTTCCGGTGA 687  
Qy 657 GCAGTGAATGCGTGAAGATATGAAGAACATCGATGCGAAGGCGACCTCTGGGTTAA 716  
Db 688 GCGGTGAATGCGTGAATATCGAAGAAACACCAATGCGAAGGCGACCTCTGGACCA 747  
Qy 717 CACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGTATCCCTGTGTGTA 776  
Db 748 TACTGACACTGAGGCGCAAGCGGTGGGAGCAACAGATTAAGACCTGTGTGTA 807  
Qy 777 CGCCCTAAAGATGCACTAGTGTGGGCTTTAAGGCTTGTGTAAGCAAGTAAAGC 836  
Db 808 CGCAGTCAACGATGGAATCTAGCTTTGGGAAGTCCCTTCTTATGTAAGCAAGTAAAGC 867  
Qy 837 GTGAAGTTGACCGCTGGGAGTACGAGTGCAGATTAACCAAGGAATGACGAGG 896  
Db 868 GTTAAGTCTCGGCTGGGAGTACGAGCGCAAGGTTAAACTCAAGAAATGACGAGG 927  
Qy 897 ACCGCAACAGCGGTGATATGTGATTAATTCATGCAACGCAAAACCTTACCTAC 956  
Db 928 GCGCCCAAGCGGTGAGCATGTGTTAATTCATGCAACGCAAAACCTTACCTAC 987  
Qy 957 CTTGACATGTAGCAATTTTCTAGAGTATGATGTG-CTTCGGGAACGCTAACACAG 1015  
Db 988 CTTGACATCTCTCGGAATCTGTCAAGATGATTTGTGCTTCGGGAACGAGTACAG 1047  
Qy 1016 TGCTGACATGCTGTGCTGCTGTGCTGTGAGATGTTGGTTAAGTCCCGCAACGAG 1075  
Db 1048 TGCTGACATGCTGTGCTGCTGTGCTGTGAGATGTTGGTTAAGTCCCGTAAACGAG 1107  
Qy 1076 CAACCTTGTCTATTATTTGCCATC--ATTGGTTGGGCACTTATGAGATGCGCGGTGA 1133  
Db 1108 CAACCTTGTCTATTATTTGCCATCAGTCAAGTCCGGAACCTTAAAGAGATGCGCGGTGA 1167  
Qy 1134 CAACCTGAGGAGGTTGGGAGTGAAGTCAAGTCTCATGAGCCCTTAAGGTTAGGCTTCA 1193  
Db 1168 TAAACCGAGGAGGTTGGGAGTGAAGTCAAGTCAATGAGCCCTTAAGGTTAGGCTTCA 1227  
Qy 1194 CACGTATATCAATGCGGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGA 1253  
Db 1228 CACGTATATCAATGCGGCTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGA 1287  
Qy 1254 AAGCGCTGTGATGCTCGGATCGGAATCTGCAACTCGATCCGTGAAGTCCGAATGCTGAG 1313  
Db 1288 AAGCGCTGTGATGCTCGGATCGGAATCTGCAACTCGATCCGTGAAGTCCGAATGCTGAG 1347  
Qy 1314 TAAATCGGATCAGATGCTCGGATGGAATGATGCTTCCGGGTTTGTAAACACCGCGCTC 1373  
Db 1348 TAAATCGGATCAGATGCTCGGATGGAATGATGCTTCCGGGTTTGTAAACACCGCGCTC 1407  
Qy 1374 ACACCATGGAGTGGGTTTCAACAGAGCAGGTACTTAAACGTAAAGAGGCGGCTTCC 1433  
Db 1408 ACACCATGGAGTGGGTTTCAACAGAGCAGGTACTTAAACGTAAAGAGGCGGCTTCAAC 1467

Qy 1434 ACGTGAGATTATCATGAC 1450  
Db 1468 ACGGTATGATCANGAC 1484  
RESULT 30  
US-09-737-297-4  
; Sequence 4, Application US/09737297  
; Patent No. 6887984  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Mark  
; APPLICANT: Griffiths, Allen  
; APPLICANT: Hill, Philip  
; APPLICANT: Laybourne-Parry, Johanna  
; APPLICANT: Mills, Sarah  
; TITLE OF INVENTION: Processes and Organisms for the Production of Antifreeze Proteins  
; FILE REFERENCE: F3247  
; CURRENT APPLICATION NUMBER: US/09/737,297  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: GB 9929696.4  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 1481  
; TYPE: DNA  
; ORGANISM: Marinomonas communis  
; FEATURE:  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (203)..(204)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (840)..(840)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (964)..(965)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (1142)..(1142)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (1182)..(1182)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (1185)..(1186)  
; OTHER INFORMATION: base identity unsure  
; NAME/KEY: Unsure  
; LOCATION: (1449)..(1449)  
; OTHER INFORMATION: base identity unsure  
US-09-737-297-4  
Query Match 69.8%; Score 1017.4; DB 3; Length 1481;  
Best Local Similarity 82.3%; Pred. No. 0;  
Matches 1196; Conservative 1; Mismatches 252; Indels 4; Gaps 3;  
Qy 1 ATGGAAGCTTGGCGGCAATGCTTTACATGCAAGTCAACGAGCAGATGCTTGCAT 60  
Db 29 ATTGAACGCTGGCGGCAAGGCTTAAACATGCAATGCAATGCAATGCTTACCTGCT 88  
Qy 61 CT--GGTGGCGAGTGGCGGAGCGGTGATGATGATGCAATGGAAACGTATCCAGAAAGGGGG 118  
Db 89 AGAAGATGACGAGCGGCGGAGCGGTGATGATGATGCAATGGAAACGTATCCAGAAAGGGGG 148  
Qy 119 TAAAGCATGGAAGATGCTAATACCGCATATCTCTAAGGAGAAAGCAGGGGATCGA 178  
Db 149 CAACATGTGGAAGCGCATGCTAATACCGCATACCCCTTACGGGGGAAAGAGGAGGATCTT 208  
Qy 179 AAGACCTTGGCTTTTGAAGCGGCGAGTGTCTGATTAAGTATGATGAGGCTTAAAGGCTT 238

Db 209 CGGACCTTTCGCTAATTAGATGAGCCTCGGTGAGATTAGCTAGTGTGGGGTAAAGGCTT 268  
Qy 229 ACCAAGGCGAGCATGAGTGTGTCTGAGAGAGACGACCGACACTGGGACTGAGACA 298  
Db 269 ACCAAGGCGAGCATGAGTGTGTCTGAGAGAGATGATCAGCACACTGGGACTGAGACA 328  
Qy 299 CGGCGCAGACTCTCTACGGGAGGCGACAGTGGGGAAATTTTGACAAATGGCGCAGCTCTGA 358  
Db 329 CGGCGCAGACTCTCTACGGGAGGCGACAGTGGGGAAATTTTGACAAATGGCGCAGCTCTGA 388  
Qy 359 TCCAGCAATGCCCGGTGAGTGAAGAGGCTTCGGGTTGTAAAGCTCTTTCACTGAGAGA 418  
Db 389 TCCAGCAATGCCCGGTGAGTGAAGAGGCTTCGGGTTGTAAAGCACTTTCAGAGGTGAG 448  
Qy 419 GAAAAGGTAGCGTAAATTAATCGTGACTGATGACGGTATCGACAGAGAGACCGGCTA 478  
Db 449 GAAAAGGTAGTATGTTAATCTCTGTATGTTTGAAGCTTCAAGAGAGAGACCGGCTA 508  
Qy 479 ACTAGGTGCGACAGCGCGGTAAATCGTAGGGTGCAGCGTTAATCGGAATTAATCTGGGC 538  
Db 509 ACTGTGTGCACAGCGCGGTAAATCGTAGGGTGCAGCGTTAATCGGAATTAATCTGGGC 568  
Qy 539 GTAAAGGGTGGCGAGCGGCTTTGTATGATCAAGTGTGAATCCCGGGCTTAACTTGGGA 598  
Db 569 GTAAAGCGCGGTAGCGGCTTTGTATGATCGAGTGTGAATCCCGGGCTTCAACTTGGGA 628  
Qy 599 ATTGGGTTGAATCTACAGGCTAGAGTGTGCGAGGAGGAGTGTGAATTCATGTTGAC 658  
Db 629 ATTGGCACCGGATCTGGCAGGCTAGAGTGTGAGGAGGAGTGTGAATTCCTGTGTAGC 688  
Qy 659 AGTGAATGCGTAGAGATATGAGAGAACATCGATGGCGAAGCGCTCTGGGTTTACA 718  
Db 689 GGTGAATGCGTAGAGATATGAGAGAACATCGATGGCGAAGCGCACCTGGACCGATA 748  
Qy 719 CTGAGCGCTATGACGAAAGCGTGGGAGCAACAGAGTTAATACCTGTGTGTCCAG 778  
Db 749 CTGAGCGCTATGAGGCGAAAGCGTGGGAGCAACAGAGTTAATACCTGTGTGTCCAG 808  
Qy 779 CCTTAAAGCATGTCATGATGTTGTGGGCTTATTTAGGCTGTGTAAGAGCTTACGCGT 838  
Db 809 CGTAAACCATGTCATGATGTTGTGGGCTTATTTAGGCTGTGTAAGAGCTTACGCGT 868  
Qy 839 GAATGACCGGCTGGGAGTACGCTCGAAGATTAAATCTCAAGAGATTGA CGGAGAC 898  
Db 869 TAAAGTACCGGCTGGGAGTACGCGCGAAGTTAAATCTCAATGAAATGACCGGAGC 928  
Qy 899 CCGCACAAGCGGTGATTTATGTGATTAATTCATGCAACGCAAAAACCTTACCTTACC 958  
Db 929 CCGCACAAGCGGTGATGTTGTTAATTCGANNAAACGCAAGAACCTTACCTTACC 988  
Qy 959 TTGACATGTAAGCAATTTTCTAGAGATAGTTAGTGTGCGGAGCGCTTACACAGGTG 1017  
Db 989 TTGACATGTAAGCAATTTTCTAGAGATAGTTAGTGTGCGGAACTCTGAGACAGGTG 1048  
Qy 1018 CTGCAATGCTGTGCTGAGCTGTGTGAGATGTTGGTTAAGTCCCGCAAGAGCGCA 1077  
Db 1049 CTGCAATGCTGTGCTGAGCTGTGTGAGATGTTGGTTAAGTCCCGCAAGAGCGCA 1108  
Qy 1078 ACCCTGTCTAATTAATGTCATCA - TTTGTTGGGCACTTTAATGAGACTGCGGTGACA 1136  
Db 1109 ACCCTGTCTAATTAATGTCATCACTTGGGTGANAACCTTCAAGAGACTGCGGTGACA 1168  
Qy 1137 ACCGAGAGAGGTGGGAGTACGTCAGTCTGATGCGCTTATGGGTAGAGGCTTACAC 1196  
Db 1169 ACCGAGAGAGGTGGGAGTACGTCAGTCTGATGCGCTTATGAGATAGAGGCTTACAC 1228  
Qy 1197 GTAAATCAATGAGCGCTACAGAGGTTGCCAACCGCGAGAGGAGCTTAACTCAGAAAG 1256  
Db 1229 GTGCTACAAATGGCTATACAGAGGCGAGAACTCGGAGAGGTAAAGCAAAATCCAAAAG 1288  
Qy 1257 CGGCTGTAGTCCGAGTCCGAGTCTGCACTCGACTCGTGAAGTCCGAATCCGTAGTAA 1316  
Db 1289 TACGTGTAGTCCGAGTCTGCACTCGACTCGTGAAGTCCGAATCCGTAGTAA 1348

Qy 1317 TCGCGGATCAGATGTCGCGGTGAATAGCTTCCGGGCTTGTACACACCGCGCGTACA 1376  
Db 1349 TCGTGAATCAAGATGTCAGGTGAATACGTTCCGGGCTTGTACACACCGCGCTACA 1408  
Qy 1377 CCATGGAGTGGGTTTACCAGAGAGAGTGTAACTTAAAGAGAGGCGCTTGCACAG 1436  
Db 1409 CCATGGAGTGGGTTTACCAGAGAGAGTGTAACTTAAAGAGAGGCGCTTGCACAG 1468  
Qy 1437 GTGAGATTCAATGA 1449  
Db 1469 GAGTGTCAATGA 1481  
RESULT 31  
US-08-114-695A-6  
Sequence 6, Application US/08114695A  
Patent No. 5508193  
GENERAL INFORMATION:  
APPLICANT: Mandelbaum, Raphael T.  
APPLICANT: MacKelt, Lawrence P.  
TITLE OF INVENTION: DEGRADATION OF S-TRIAZINES IN SOIL AND  
TITLE OF INVENTION: WATER  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCHWEGMAN, LUNDBERG & WOESSNER, P.A.  
STREET: 3500 IDS CENTER  
CITY: MINNEAPOLIS  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/114,695A  
FILING DATE: 31-AUG-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MUEITING, ANN M.  
REGISTRATION NUMBER: 33,977  
REFERENCE/DOCKET NUMBER: 600,268051  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-339-0331  
TELEFAX: 612-339-3061  
INFORMATION FOR SEQ. ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1518 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: rRNA  
ORIGINAL SOURCE:  
ORGANISM: Pseudomonas aeruginosa  
US-08-114-695A-6  
Query Match 69.0%; Score 1005; DB 2; Length 1518;  
Best Local Similarity 67.1%; Pred. No. 0;  
Matches 974; Conservative 235; Mismatches 237; Indels 6; Gaps 6;

```
QY 180 AGACCTTGCGCTTTTGGAGCGGCCGATGTTCTGATTAGTGAAGTGGGGTTAAAGCCTTA 239
Db 206 GGAACCCGACGCGGAGUAGAACCTUAGCGGUAUUGGUGGAGUAAAGCCUUA 265
QY 240 CCAAGCGACGATCGATGTTGGTCTGAGAGACGACGACCTGGGATCTGAGACAC 299
Db 266 CCAAGCGACGATCGGUAACUGGUCUGAGAGGACGUAUCAGUACACUGGAACTGAGACAC 325
QY 300 GGGCCGACCTCTTACGAGGAGGACGAGGAGGAGATTTTGAACATGGGCGGACCTGAT 359
Db 326 GGUCCAGACUCCUACGAGGAGGACGAGGAGGAGUUAUUGACAUAUGGGGAAAGCCNANU 385
QY 360 CCAAGCATGCGCGCTGATGAGAAAGGCTTGGGTTGTAAGGCTTTTCACTGACAGAG 419
Db 386 CCAAGCATGCGCGGUGUAGAAAGGUCUUCGAAUUGAAAGCAUUAUUGGAGG 445
QY 420 AAAAGTTACGGTAATATCGTACATGACGGGATCGACAGAAAGACCGGCTTA 479
Db 446 AAGGCGAUAAGUUAUUAUACUUGUUGUACGUUACCAACGAAUUAAGCACCGGCUAA 505
QY 480 CTACGTGCGACGACGCGCGTAAATCGTAGGGTGAAGGCTTAATCGGAATTACTGGGCG 539
Db 506 CUUCGUGCAGACGCGCGGUAUUAAGAAAGGUGGAGGUGUAUUCGGAUUAUCUGGCG 565
QY 540 TAAAGGGTCCGAGCGGCTTTGTAATGTCAGATGTAAGTCCCGGGCTTAACCTGGGAA 599
Db 566 UAAAGCGCGGUGAGUGUUCGACAGUUGUAGAAUCCCGGCGUACCAACUGGAA 625
QY 600 TTGCGTTTGAACCTCAAGGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTAGCA 659
Db 626 CUGCAUCCMAAACUACUGAGCUAGUACGUAAGGUGGUGGUAUUCUGUUAACG 685
QY 660 GTGAATGCTGATGAGATATGAAAGCAATCGATGCGAAAGCGCTTCCTGCTTAACAC 719
Db 686 GUGAAAUUCGUAUUAUAGAAAGAACACAGUGGCGAAAGCGACACUGGACUGAUA 745
QY 720 TGACGCTCATGACGAAAGCGTGGGAGCAACAGATTTAGTACCTTGATGTCACAGC 779
Db 746 UGACACUAGGUGCGAAAGCGGAGGAGCAACAGAUUAUACCGUGUAGUCACGCG 805
QY 780 CCTAAACGATGCACTAGTGTGAGGCTTATTTAG-GTTGTGAACGAAAGCTAAGCGGT 838
Db 806 CGUAAACGATGCGACUAGCGGUGGAGUCCUGAGAUUCUUAUGGCGGACUAAACGGA 865
QY 839 GAAGTTGACCGGCTTGGAGATCGGTCCGACAGATTAAACCTCAAGAAATGACGAGAC 898
Db 866 UAAAGCGACCGGCGGAGUACGCGCGCAAGUUAUUAUUAUUAUUAUUAUUAUUAUUA 925
QY 899 CCGGACAGCGGTGATATGATGATTAATTCGATGCAACCGGAAACCTTAACCTACCC 958
Db 926 NNGCAAGCGGUGAGUAGUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 985
QY 959 TTGACATGATGAGAAATTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTAACAGGTG 1017
Db 986 UUGACAUUGCUAGAACTUUCGAGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 1045
QY 1018 CTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077
Db 1046 CUGCAUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 1105
QY 1078 ACCCTTGATTTAATTTGCTATCA-TTTGGTTGGGACATTTAATGACATGCGCGGTGACA 1136
Db 1106 ACCCUUGGCUUAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 1165
QY 1137 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1196
Db 1166 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1225
QY 1197 GTAATACATGAGCGCGGTGACAGAGGTTGCAACCGCGAGGAGGAGGAGGAGGAGGAGG 1256
Db 1226 GUGGUAACAUGUGGUGGUAACAAGGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1285
```

```
QY 1257 GCGGCTGATGCGGAGTGCAGATCGAATCTCCGCTGAGTGGAGTGGATCGCTAGTAA 1316
Db 1286 CCGAUCGUAUGCCGGAUUGCAGUUCUACUACUACUACUACUACUACUACUACUACUACU 1345
QY 1317 TCGGAGTACAGATGTCGCGGTGAATAGTTCCCGGCTTTGTATACACACCGCGCTCA 1376
Db 1346 UCGUAUACAAGUACUACGUGUAUACGUGUCCGGGCGUGUACACACCGCGGUGACA 1405
QY 1377 CCAATGGAGTGGGCTTTCACGAAAGCAGTATGTTAAACCTTAAGAGAGGCGCTTGCCACG 1436
Db 1406 CCAATGGAGTGGGCTTTCACGAAAGCAGTATGTTAAACCTTAAGAGAGGCGCTTGCCACG 1465
QY 1437 GTGAGATTATG 1448
Db 1466 GAGUGAUAUG 1477

RESULT 32
US-10-007-527A-12
; Sequence 12 Application US/10007527A
; Patent No. 6949362
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kosticbka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Rhodococcus AM12
US-10-007-527A-12

Query Match 68.9%; Score 1004.4; DB 3; Length 1424;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 1171; Conservative 0; Mismatches 221; Indels 4; Gaps 3;

QY 65 TGGCGATGTCGCGGACGCGGTGATGATCGGAACGATTCAGAAAGAGGCGGTAAAGC 124
Db 29 TTGAGAGCGCGGACGCGGTGATGATCGGAATCGCTTACGGAAGAGGCGGTAAAGC 88
QY 125 ATCGAAAGATGCTAATATCCGATATCTTAAGAGAGGAGGAGGAGGATCGAAAGAC 184
Db 89 TCGAAAGAGACGCTAATATCCGATATCGCTTACGGAAGAGGAGGAGGAGGAGGAGGAGG 148
QY 185 TTGCGCTTTTGGAGCGCGGATGCTGATTTAGTTGGTGGGTAAAGGCTTACCAAG 244
Db 149 TTGCGCTATCAGATGAGGCTTGTGATTTAGTTGGTGGGTAAAGGCTTACCAAG 208
QY 245 GCGACATCACTAGTTGCTGAGAGACGACGACCACTGGGACTGAGACACGCGCC 304
Db 209 GCGACATCACTAGTTGCTGAGAGATGATCACTGAGACGAGGAGGAGGAGGAGGAGGAGG 268
QY 305 AGACTCTTACGAGGAGGACGAGTGGGATTTTGAACAATGGGCGCAAGCCTGATCCAGC 364
Db 269 AGACTCTTACGAGGAGGACGAGTGGGATTTTGAACAATGGGCGCAAGCCTGATCCAGC 328
QY 365 AATGCGCGGTGAGTGAAGAAAGCCTTGGGTTGTAAGCTTTCACTGAGAAAGAAAG 424
Db 329 CATGCGCGGTGAGTGAAGAAAGCCTTGGGTTGTAAGCTTTCACTGAGAAAGAAAG 388
QY 425 GTTACGCTAATTAATCTGATCACTAGACGATGATGACAGAAAGAAAGCAGGCTTAAC 484
Db 389 CAGTTACCTAATTAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 448
QY 485 TGCAGAGACCGCGGTTAATCGTAGGGTGAAGGCTTAATCGAATTAATCGGCGTAAAG 544
```

Db ||||| 449 TCCCGACGCCCCGGTAATACAGAGGGTCCAGCGTTAAATCGAATATTCTGGCGCTAAAG 508  
Qy ||||| 545 GGTGCGCAGCGCGCTTTGTAAGTCAGATGTGAATCCCGCGGCTTAACCTGGGAATTTGCG 604  
Db ||||| 509 CCGCGCTAGGTGTTGTTAGTGTGAATGTGAATCCCGGGGTCAACCTGGGAATTTGCA 568  
Qy ||||| 605 TTTGAAACTACAGAGCTAGAGTGTGGCAGAGGGAGGTGAATTCATGTGTAGCACTGAA 664  
Db ||||| 569 TTCAAACTGACTGACTAGAGTATGTGTAGAGGGTGTGTAATTTCTGTGTAGCGGTGAA 628  
Qy ||||| 665 ATGCGTAGAGATGTGAAGAACATCGATGGCCGAAGGACGCTTCCTGGGTTAACACTGAG 724  
Db ||||| 629 ATGCGTAGATGTGAAGAACACAGTGGCCGAAGGACGACCACTGAGACTGATGATGACA 688  
Qy ||||| 725 CTCATGACGAAAGCGTGGGGAGCAAAAGAGATTATATACCTCGTAGTCCAGCCCTAA 784  
Db ||||| 689 CTGAGGTGCGAAGCGTTGGGGAGCAAAAGAGATTATATACCTCGTAGTCCAGCCCTAA 748  
Qy ||||| 785 AGCATGTCAACTAGTGTGTGGGCTTATTAAG-CTTGGTAACGAGCTAACCGGTGAAGT 843  
Db ||||| 749 AGCATGTCAACTAGCGGTTGGGAGCCTTAGCTCTTAGTGGCGCAGCTAACCGCATTAAGT 808  
Qy ||||| 844 TGAACGCTTGGGGAGTACCGTCCGAAGATTAAATCTAAAGATTGAAGGGGACCCGCA 903  
Db ||||| 809 TGACCCGCTGGGGAGTACCGCCGCAAGTTAAATCTAAATGAATTGAAGGGGGCCCGCA 868  
Qy ||||| 904 CAAGCGGTGGATTATGTGAATTAAATTCAGATGCAACGGAAGAAACCTTACCTACCTTTCAC 963  
Db ||||| 869 CAAGCGGTGGAGATGTGTTTAATTCAGAGCAACGGAAGAAACCTTACCAAGGCTTTGAC 928  
Qy ||||| 964 ATGTACGAAATTTTCTAGAGATGATTAAGTG-CTTGGGAAACGCTAACAGAGTGTCTGA 1022  
Db ||||| 929 ATGTACGAAATTTTCTAGAGATGATTAAGTGCTTGGGAAACATTTGAGACAGGTGTCTGA 988  
Qy ||||| 1023 TGGCTGTCTGACGTGTGTCTGAGATGTGGGTTAAGTCCCGCAACAGGCGCAACCTT 1082  
Db ||||| 989 TGGCTGTCTGACGTGTGTGTGAGATGTGGGTTAAGTCCCGTAAACAGGCGCAACCTT 1048  
Qy ||||| 1083 TGTGATTAATTTGCTATCA--TTTGTGGGCACTTAAATGAGACTGCGCGGTGACAAACG 1140  
Db ||||| 1049 TGTGCTTAAGTTACAGACCTTAATGTGGGCACTTAAGAGACTGCGCGGTGACAAACG 1108  
Qy ||||| 1141 GAGGAAGTGGGGATGACGTCAAGTCTCTCATGGCCCTTAATGGGTAAGGCTTTCACAGTAA 1200  
Db ||||| 1109 GAGGAAGTGGGGATGACGTCAAGTCTATATGGCCCTTAAGGCTTTCACAGAGTGC 1168  
Qy ||||| 1201 TACAAATGGCGGTACAGAGGGTTCGCAACCGCGAGGGGAGCTAATCTCAAGAAAGCGG 1260  
Db ||||| 1169 TACAAATGGTGTGACAGAGGGTTCGCAACCGCGAGGGTGAAGCTAATCCAGAAACCGA 1228  
Qy ||||| 1261 TCGTATTCGCGAATCGAGTCTGCACTCGAATCGGAAATCGGTAGTAATCGC 1320  
Db ||||| 1229 TCGTATTCGCGAATCGAGTCTGCACTCGAATCGGAAATCGGTAGTAATCGC 1288  
Qy ||||| 1321 GATTCAGCATGTGCGCGGTGAATAGTCCCGGGCTTGTACACACCGCGCTGACACCAT 1380  
Db ||||| 1289 GATTCAGCATGTGCGCGGTGAATAGTCCCGGGCTTGTACACACCGCGCTGACACCAT 1348  
Qy ||||| 1381 GGGAGTGGGTTTCAACGAAGAGGTATGTTAAACCGTAAGAGAGGGCGCTTGCACGGTGA 1440  
Db ||||| 1349 GGGAGTGGGTTTCAACGAAGAGGTATGTTAAACCGTGGGAGGAGCGGTTTACACGGTGT 1408  
Qy ||||| 1441 GATTTCAGACTGGGCT 1456  
Db ||||| 1409 GATTTCAGACTGGGCT 1424

RESULT 33  
US-09-726-774-14  
; Sequence 14, Application US/09726774  
; Patent No. 6677153  
; GENERAL INFORMATION:

; APPLICANT: Iversen, Patrick L.  
; TITLE OF INVENTION: Antisense Antibacterial Method and  
; FILE OF INVENTION: Composition  
; FILE REFERENCE: 0450-0032.30  
; CURRENT APPLICATION NUMBER: US/09/726,774  
; CURRENT FILING DATE: 2000-11-29  
; PRIOR APPLICATION NUMBER: US 60/168,150  
; PRIOR FILING DATE: 1998-11-29  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 1487  
; TYPE: DNA  
; ORGANISM: Shigella dysenteriae  
US-09-726-774-14  
Query Match 68.5%; Score 998.2; DB 3; Length 1487;  
Best Local Similarity 82.0%; Pred. No. 0;  
Matches 1199; Conservative 0; Mismatches 258; Indels 6; Gaps 4;  
Qy ||||| 1 ATTGAACGCTGCGCGCATGCTTTTACATGCAAGTGCAGCG--CAGCAGCATGCTTG 57  
Db ||||| 9 ATTGAACGCTGCGCGCATGCTTTTACATGCAAGTGCAGCGTAAACAGAAAGCAGCTTGC 68  
Qy ||||| 58 CATCTGTGGCGAGTGGCGGAGCGGTGATATGCAATCGGAACGTAACGAAAGAGGGGG 117  
Db ||||| 69 TGTGTCTACAGATGGCGGAGCGGTGATATGCTGGAATCTGCTGATGAGAGGG 128  
Qy ||||| 118 GTACGATCGAAGAAAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGCAGGGATG 177  
Db ||||| 129 ATTAATCTGGAAGAAAGATGTGCTAATACCGCATATCTGCAAGAACAAAGAGGGAGCT 188  
Qy ||||| 178 AAAGACCTTGCCTTTTGAAGCGCGCATGCTGATTAAGTGTGTGTGGTGAAGGCC 227  
Db ||||| 189 TCGGCGCTCTTGCATCGGATGATGCTCAAGATGGGATTAAGTGTGGGTGAAGCGCT 248  
Qy ||||| 238 TACCAAGGAGAGATGATGATGCTGAGAGAGACACAGCACTGAGGAGCTAGAG 297  
Db ||||| 249 CACTAGGCGAGCATCTTACGCTGTGTGAGAGATGACAGCCACACTGGAACCTAGAG 308  
Qy ||||| 298 ACGGCGCAGATCTCTACGCGGAGCAGCAGTGGGAAATTTTGAACAATGGCGCAGCCTG 357  
Db ||||| 309 ACGGCGCAGATCTCTACGCGGAGCAGCAGTGGGAAATTTTGAACAATGGCGCAGCCTG 368  
Qy ||||| 358 ATTCAGCAATGCGCGGTGATGAGAAAGGCTTGGGTTGAAGCTTTTCTGATGAGA 417  
Db ||||| 369 ATTCAGCAATGCGCGGTGATGAGAAAGGCTTGGGTTGAAGCTTTTCTGATGAGA 428  
Qy ||||| 418 AGAAAGTTTACGTTAATATGATGATCAAGACGTTATCCAGAGAAAGCAGCGGCT 477  
Db ||||| 429 GGAAGGAGTTAAGTTAATATGATGATCAAGACGTTATCCAGAGAAAGCAGCGGCT 488  
Qy ||||| 478 AACTAGTGCAGCAGCGCGGTATATAGTAAAGGTTCAAGCGTTAATCGGAATTAAGT 537  
Db ||||| 489 AACTAGTGCAGCAGCGCGGTATATAGTAAAGGTTCAAGCGTTAATCGGAATTAAGT 548  
Qy ||||| 538 CGTAAAGGTTGCGCAGCGGCTTTTGAAGTCAAGTGAATCCCGGCTTAACTGAGG 597  
Db ||||| 549 CGTAAAGGTTGCGCAGCGGCTTTTGAAGTCAAGTGAATCCCGGCTTAACTGAGG 608  
Qy ||||| 598 AATTGCTTTGAATCAAGGTTAGAGTGTGCAAGAGGAGTGAATTCATGCTAG 657  
Db ||||| 609 AACTGATCTGATATGCGCAAGCTTATGCTCTTAAGAGGGGGTGAATTCAGAGTGA 668  
Qy ||||| 658 CAGTGAATGCGTGAATATGGAAGAACATGATGAGCGAAGGACGCTCTGAGTTAAC 717  
Db ||||| 669 CAGTGAATGCGTGAATATGGAAGAACATGATGAGCGAAGGACGCTCTGAGTTAAC 728  
Qy ||||| 718 ACTGACGCTCATGACGAAGAGCGTGGGAGCAAAAGATTTAGATACCTGTGATGCTCAC 777  
Db ||||| 729 ACTGACGCTCATGAGTGAAGAGCGTGGGAGCAAAAGATTTAGATACCTGTGATGCTCAC 788  
Qy ||||| 778 GCCCTAAAGATGATCAACTAGTGTGGGCTTATTAAGGCTTGG-TAACGAAGCTTAAGC 826



```
Db 789 GCCGTAACGATGTCGACTTGGAGGTTGGCCCTTGAAGCGGTGGCTTCGGAGCTAAAGC 848
Qy 837 GTGAAATTGACCGCTGGGAGTACGTCGCAAGATTAAACTCAAGGAATTGACGGGG 896
Db 849 GTTAAGTCGACCGCTGGGAGTACGCGCCGCAAGGTTAAACTCAAAAGAAATTGACGGGG 908
Qy 897 ACCGCAACAGCGGAGATTAAATGTGAAATTAAATTCGATGCAACGCAAAAACCTTACCTAC 956
Db 909 GCCGCAACAGCGGAGTACGATGTGGTTAAATTCGATGCAACGCAAAAACCTTACCTAC 968
Qy 957 CTTTGAACATGATCGCAATTTTCTAGAGATAGATTAGTG-CTTCGGGAAACGCTAACACAG 1015
Db 969 TCTTACATCCACAGAACCTTGTAGAGATACAGAGGTGCTTCGGGAACTGTAGACAGG 1028
Qy 1016 TCTGCAATGCTGTGTCAGCTGTGTGTCGAGATGTGGTTAAATCCCGCAACGACG 1075
Db 1029 TGCTGCAATGCTGTGTCAGCTGTGTGTCGAAATGTTGGTTAAGTCCCGCAACGACG 1088
Qy 1076 CAACCTTGTGCTTAAATTGCGCATC-AITTTGTTGGGCACTTAAATGAGACTGCGGTAC 1134
Db 1089 CAACCTTATCTTGTGTCGACGCGTCCGCGGAACTCAAGAGACTGCGCATGAT 1148
Qy 1135 AAACCGAGAGAGGTGGGATGACGTCAAGTCTCATGCGCTTATGGGTAGGCTTAC 1194
Db 1149 AAACCGAGAGAGGTGGGATGACGTCAAGTCAATGCTCATGCGCTTATGACAGGGCTTAC 1208
Qy 1195 ACGTAAATCAATGCGCGCTACAGAGGTTGCCAACCCCGAGAGGGAGCTTAATCTCAAA 1254
Db 1209 ACGTCTCAATATGGCGCATACAAAGAGAGCACTCCGAGAGAGAGCGGACCTCATTA 1268
Qy 1255 AGCGGTGTGTCGCGATCGGATCGGAATCTGCAACTCGACTCCGTAAGTGGAAATGCTAGT 1314
Db 1269 AGTGTGTGATGTCGGAATGGAGTCTGCAACTCGACTCGACTGAAGTGGAAATGCTAGT 1328
Qy 1315 AATCGCGATGACGATGTCGCGGTGAATACGTTCCCGGCTCTTGTACAACCGCCGTCA 1374
Db 1329 AATGTGTGATGAGATGTCAGGTGATACGTTCCCGGCTTGTACAACCGCCGTCA 1388
Qy 1375 CACCATGGAGAGTGGTTTCAACCAAGACAGTATCTTAACCGTAAGAGGGGCTTTGCCA 1434
Db 1389 CACCATGGAGAGTGGTTTCAAAAGAGATAGTACTTAACCTTCGAGGGGCTTTACCA 1448
Qy 1435 CGGTGAGATTGATGACTGGGGTG 1457
Db 1449 CTTGTGATCATGACTGGGGTG 1471

RESULT 34
US-09-726-774-3
; Sequence 3, Application US/09726774
; Patent No. 6677153
; GENERAL INFORMATION:
; APPLICANT: Iversen, Patrick L.
; TITLE OF INVENTION: Antisense Antibacterial Method and
; TITLE OF INVENTION: Composition
; FILE REFERENCE: 0450-0032.30
; CURRENT APPLICATION NUMBER: US/09/726,774
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 60/168,150
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-726-774-3

Query Match 68.4%; Score 997.2; DB 3; Length 1467;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 1165; Conservative 0; Mismatches 223; Indels 4; Gaps 3;
```

```
Qy 70 AGTGGCGACGGGTGATGATGATCGGACGTATCCAGAAAGAGGGGGTAAACCATGCA 129
Db 26 AGCGCGGAGCGGTGATGATGATGCTAGGAAATCTGCTGATAGTGGGGGCAACCGTTTGC 85
Qy 130 AAGATGTGCTAATACCGCATATCTTAAGAGAGAAAGACAGGGGATCGAAAGACTTTGCG 189
Db 86 AAGGAAAGCGTAAATCCGATACGTCTAAGGGAAGAAAGACAGGGGACCTTCGGGCTTGGC 145
Qy 190 CTTTGGAGCGGCGGATGTGCTGATTTAGCTAGTTGGGGTAAAGGCTTACCAAGCGCAC 249
Db 146 CTATGATGATGACCTTAGGTGCGATTTAGCTTAGTTGGTGAAGTTACGCTCACCAAGGCGAC 205
Qy 250 GATCAGATGATGCTCTGAGAGAGACGACGACCACTGGGACTAGACACGCGCCAGACT 309
Db 206 GATCGTACATGCTGTGAGAGAGATGATCACTGACACTGAGACTAGACAGGCTCAGACT 265
Qy 310 CTTACGAGAGGACAGCATGCGGAAATTTTGAACAATGGGCGCAGCTGATCCAGCATGC 369
Db 266 CTTACGAGAGGACAGCATGCGGAAATTTTGAACAATGGGCGAAGCTGATCCAGCATGC 325
Qy 370 CGCGTGAAGTGAAGAGCGCTTCGGGTTGAAGCTCTTTCAGTGAAGAAAGGTTAC 429
Db 326 CGCGTGTGAAGAGCGCTTCGGAATGTGAAGCACTTAAATGTTGGAGAGAGGCAATTA 385
Qy 430 GGTAAATATGTCGATCATGACGTATCGACAGAAAGACCGGCTTAAGTACGTGCA 489
Db 386 AACTTAATATGTTAGTGTGTTACCGTTACGACAGAAATAGCACCGGCTTACTTGTGCA 445
Qy 490 GCAGCGCGGTAAATCGTAAAGGTGCAAGCTTAAATCGAAATTAATCTGGGCTTAAAGGTC 549
Db 446 GCAGCGCGGTAAATCGAAAGGTGCAAGCTTAAATCGAAATTAATCTGGGCTTAAAGGTC 505
Qy 550 GCAGCGCGCTTGTAAAGTCAATGTAATCCCGGCTTAACTGAGGAAATTCGCTTGA 609
Db 506 GTAGGTGTTGTTAATGTTAAATGTGAAGCCCGGCTCAACTGGGAACTGCATCAAA 565
Qy 610 AACTTAAGGCTAGAGTGTGCAAGAGGAGTGAATTCATGTGTAGCACTGAATTCG 669
Db 566 AACTGGAAGCTAGAGTGTGCAAGAGGAGTGAATTCCTGTGTAGCGGTGAATTCG 625
Qy 670 TAGAGATTTGAAGAAATCATGATGGCGAAGGCAAGCTTCTGGGTTAACTGAAGCTCAT 729
Db 626 TAGATATGAAGAAACCAAGTGGCGAAGGCAACCTGGGTTAAATCACTGAACCTGAG 685
Qy 730 GCAGAAAGCGTGGGAGCAAAACAGATTAGATCCCTGTATGTCACGCCCTAAACGAT 789
Db 686 GTGCAAAAGCGTGGGAGCAAAACAGATTAGATCCCTGTATGTCACGCCCTAAACGAT 745
Qy 790 GTCAACTAGTGTGGGCTTATTTAGG-CTTGTGAACGAACTTAACGGGTGAAGTTGAC 848
Db 746 GTGCACTAGCGGTTGGGATCCTTGAGATCTTAATGTGGGCACTTAACGGATTAAGTCGAC 805
Qy 849 GCTTGGGAGTACCGTGGCAAGATTAAATCTCAAAAGAAATTGAGAGGGGACCCGCAAGC 908
Db 806 GCTTGGGAGTACCGGCTGAGTTAAATCTTAATGAAATGAGAGGGGACCCGCAAGC 865
Qy 909 GTGTGATTATGATGATTAAATTCGATGCAAGCGGAAAACTTACCTACCTTGACATGTA 968
Db 866 GTGTGATGATGATTAAATTCGATGCAAGCGGAAAACTTACCTACCTTGACATGTA 925
Qy 969 GCGAATTTCTAGAGATAGATTAGTG-CTTGGGAAACGCTTAACACAGGTGCTGACGCT 1027
Db 926 GAGAACTTTCAAGAGATGAGTTGGGCTTGGGAACTCTGACACAGGTGCTGACGCT 985
Qy 1028 GTGCTACGCTGTGTGTCGATGATGTTGGGTTAAAGTCCCGCAACGAGCCCACTTGTCA 1087
Db 986 GTGCTACGCTGTGTGTCGATGATGTTGGGTTAAAGTCCCGTAAACGAGCCCACTTGTGC 1045
Qy 1088 TTAATGTCATCA--TTTGTGGGCACTTAAATGAGACGCGGTGCAAAACGAGAGA 1145
Db 1046 TTAGTTACACAGATTAAAGGTGGCACTCTTAAGAGACTCCGTTGACAAACGAGAGA 1105
Qy 1146 AGTGGGAGTACGTCAAGTCCATGCGCTTATGGGTAGGGCTTCAACGTAATACAA 1205
```



Db 1106 AGGTGGGGATGACGTCAGATCATATGCGCCCTTACGGCCCTGGCTACACACGTCCTACAA 1165  
 Qy 1206 TGGCCGCTACAGAGGCTTCCCAACCCCGAGGGGAGCTAATCTAGAAAGCGGCTGTA 1265  
 Db 1166 TGGTGGGTAACAAGGGGTGGCCAAAGCCGAGGTGAGCTAATCCATTAACCAATCGTA 1225  
 Qy 1266 GTCCGGATCGAGTCTGCAACTGCACTCCGTAAGTCGGAATCGTAGTAATCGCGGATC 1325  
 Db 1226 GTCCGATCGAGTCTGCAACTGCACTCCGTAAGTCGGAATCGTAGTAATCGCGGATC 1285  
 Qy 1326 AGCATGTCCGCTGTAATAGCTTCCGAGTCTTGTACACACCCCGCTCACACCATGAGAG 1385  
 Db 1286 AGAATGTCAACGGTGAATAGCTTCCGAGGCTTGTACACACCCCGCTCACACCATGAGAG 1345  
 Qy 1386 TGGGTTTCCACCAAGAGGCTGTCTAACCCTAAGGAGGCGCTTCCGAGGTGAGATTC 1445  
 Db 1346 TGGGTTTCCACCAAGAGGCTGTCTAACCCTAAGGAGGCGCTTCCGAGGTGAGATTC 1405  
 Qy 1446 ATGACTGGGGTG 1457  
 Db 1406 ATGACTGGGGTG 1417

RESULT 35

US-08-757-653-158

Sequence 158, Application US/08757653

Patent No. 5843669

GENERAL INFORMATION:

APPLICANT: Kaiser, Michael W.

APPLICANT: Lyamichev, Victor I.

APPLICANT: Lyamichev, Natscha

TITLE OF INVENTION: Cleavage Of Nucleic Acid Using

TITLE OF INVENTION: Thermostable FEN-1 Endonucleases

NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESSES:

ADDRESSES: Medlen & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States Of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/757,653

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: FORS-02565

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 158:

SEQUENCE CHARACTERISTICS:

LENGTH: 1542 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-757-653-158

Query Match

Best Local Similarity 68.3%; Score 995.2; DB 2; Length 1542;

Matches 1204; Conservativity 0; Mismatches 253; Indels 7; Gaps 5;

1 ATTGAACGCTGGCGGAGCTTTTACACATGCAAGTCCGACGACGAG--GATGCTTGC 58

Db 28 ATTGAACGCTGGCGGAGGCGCTTAACATGCAAGTCCGACGAGTAACAGAAAGCTTGC 87  
 Qy 59 ATC--TGTGCGAGGTGCGGACCGGTGAGTAATGATCGGAACGTAATCCAGAAAGGGG 116  
 Db 88 TTCTTGTGTCAGAGATGCGGACCGGTGAGTAATGATCGGAACGTAATCCAGAAAGGGG 147  
 Qy 117 GGTAAACGATGGAAGATGCTTAATACCGATTAATCTAAGAGGAAAGAGGGGATC 176  
 Db 148 GATTAATCTAGAAACGGTAACTAATACCGATTAATCTAAGAGGAAAGAGGGGATC 207  
 Qy 177 GAAAGACCTTGCGCTTTTGGAGCGCGCATGATCTGATTAAGTATGTTGGGTAAAGGC 236  
 Db 208 TTCCGGCCCTTTGTCATGGAATGTCGCAAGATGGAATTAAGTAAAGTGGGTAAAGGC 267  
 Qy 237 CTACCAAGGCGAGATCAATGATGTTGTTGAGAGGACCAAGCCACTGGAGATTGGA 296  
 Db 268 TCACCTAGGCGACGATCCCTAGCTGTGTGAGAGATACCAAGCCACTGGAATCTGGA 327  
 Qy 297 CACGCGCCGAGCTCTTAAGGAGGCGAGGAGTGGGAAATTTTGAACAATGGCGCAAGCT 356  
 Db 328 CACGGTCCAGACTCTTACGAGGAGGCGAGGAGTGGGAAATTTTGAACAATGGCGCAAGCT 387  
 Qy 357 GATCCAGCAATGCGCGTGAAGTGAAGAGGCGCTTGGGTTGTAAAGCTTTCACTGAG 416  
 Db 388 GATGAGCGATGCGCGCTGTATGAAAGAGGCGCTTGGGTTGTAAAGCTTTCACTGAG 447  
 Qy 417 AAAAAAGGTTACGGTAAATTAATCGTACTATGACGATATGACAGAAAGACCGGC 476  
 Db 448 AGGAAGGAGTAAAGTAAATTAATCGTACTATGACGATATGACAGAAAGACCGGC 507  
 Qy 477 TAACATGCTGCAAGAGCGCGGTAAATCTGAGGTGCAAGGCTTAATCGAATTAATCTG 536  
 Db 508 TAACATGCTGCAAGAGCGCGGTAAATCTGAGGTGCAAGGCTTAATCGAATTAATCTG 567  
 Qy 537 GCGTAAAGGAGTGGCGAGGCGCTTGTAAAGTGAATGTAATCCCGGCTTAATCTG 596  
 Db 568 GCGTAAAGGAGTGGCGAGGCGCTTGTAAAGTGAATGTAATCCCGGCTTAATCTG 627  
 Qy 597 GAATGCGCTTGAATTAATGAGTGTGCGAGGAGGAGTGGATTTCCATGTA 656  
 Db 628 GAATGCGCTTGAATTAATGAGTGTGCGAGGAGGAGTGGATTTCCATGTA 687  
 Qy 657 GCAATGAAATGCGTGAAGTATGGAAGACATGAGTGGAGGAGGAGGAGGAGGAGGAGG 716  
 Db 688 GCGTAAAGTGGTGAAGTATGGAAGACATGAGTGGAGGAGGAGGAGGAGGAGGAGG 747  
 Qy 717 CACTGACGCTACGAGTGAAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 776  
 Db 748 GACTGACGCTACGAGTGAAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 807  
 Qy 777 CGCCTAAAGCATGTCATGATGTTGGGCTTATTAAGGCTTGG--TAAAGAGCTTAACG 835  
 Db 808 CGCCTAAAGCATGTCATGATGTTGGGCTTATTAAGGCTTGG--TAAAGAGCTTAACG 867  
 Qy 836 CGTGAAGTGAACCGCTTGGGAGTACGTCGCAAGTATTAATCTCAAGAAATTTGACGG 895  
 Db 868 CGTTAAGTGAACCGCTTGGGAGTACGTCGCAAGTATTAATCTCAAGAAATTTGACGG 927  
 Qy 896 GACCGGCAAGAGGAGTGAATGATGATTAATTCGATGCAAGCGCAAAACCTTAACCT 955  
 Db 928 GACCGGCAAGAGGAGTGAATGATGATTAATTCGATGCAAGCGCAAAACCTTAACCT 987  
 Qy 956 CCCTGACATGATGCAATTTTCTAGAT--AGATTAGTCTTGGGAGCGCTTAACACAG 1014  
 Db 988 GTCTTGAATCCAGGAGATTTTCAAGATGAGAAATGTGCTTCCGGAACCGTAGACAG 1047  
 Qy 1015 GTCTGATGATGCTGTGTCAGCTGCTGTGAGATGTTGGTAAATCCCGGAAGAGC 1074  
 Db 1048 GTCTGATGATGCTGTGTCAGCTGCTGTGAGATGTTGGTAAATCCCGGAAGAGC 1107  
 Qy 1075 GCAACCTGTCAATTAATGCGATC--ATTGTTGGGACCTTAATGAGATCCCGGTGA 1133  
 Db 1108 GCAACCTGTCAATTAATGCGATC--ATTGTTGGGACCTTAATGAGATCCCGGTGA 1167

QY	1134	CAAAACCGGAGGAAGGTGGGGGATGACGTCAAGTCTCAATGCGCCCTTAATGGATAGGGCTTCA	1139
Db	1168	TAAACTGAGGAGGAAGGTGGGGATGACGTCAAGTCATCAATGCGCCCTTAACACACAGGGCTTCA	1227
QY	1194	CACGTAATATACAATGGCGCGTACAGAGGGTGGCAACCGCGAGGGGGAGCTAATCTCAGA	1253
Db	1228	CACGGCTTACATGGCGCATACAAAGAGAGCGACCTTGGAGACAAAGCGGACTTCAAT	128
QY	1254	AAGCGCTGCTAGTTCGCGATCGGAGTCTGCAACTGACTCGGTGAAGTCGGAATGCTTAG	1313
Db	1288	AAGTGCCTGCTAGTTCGCGATTCGGAGTCTGCAACTGACTCATAAGATCGGAATGCTTAG	1347
QY	1314	TAATTCGCGGATCAGCAGTCTCGCGGTGAATACGTTCCGGGCTTTGTACACACCGCCGCTC	1373
Db	1348	TAATTCGTGATCAGAAATGCCACGCGGAAATACGTTCCGGGCTTTGTACACACCGCCGCTC	1407
QY	1374	ACACCATGGAGTGGGTTTCAACAGAGCAGGTAGTCTAACCGTAAAGAGGGCGCTTCC	1433
Db	1408	ACACCATGGAGTGGGTTTCAAAAGAAAGTAGGTACTTAACCTTTCGGAGGGCGCTTACC	1467
QY	1434	AACGTGAGATCATGATCGGGGTG	1457
Db	1468	ACTTTGATTCATGATCGGGGTG	1491

```

RESULT 36
US-09-465-355-2
/ Sequence 2, Application US/09465355
/ Patent No. 6316194
/ GENERAL INFORMATION:
/ APPLICANT: Karn, Jonathan
/ APPLICANT: Knowles, David
/ APPLICANT: Murchie, Alastair
/ APPLICANT: Lenzzen, Georg
/ TITLE OF INVENTION: Methods and kits for Discovery of RNA-Binding Antimicrobials
/ FILE REFERENCE: 22650/1150 (Formerly 3950/85276)
/ CURRENT APPLICATION NUMBER: US/09/465,355
/ CURRENT FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: US 09/325,601
/ PRIOR FILING DATE: 1999-06-03
/ PRIOR APPLICATION NUMBER: GB 9812196.5
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: GB 9904790.4
/ PRIOR FILING DATE: 1999-03-02
/ PRIOR APPLICATION NUMBER: US 60/122,439
/ PRIOR FILING DATE: 1999-03-02
/ PRIOR APPLICATION NUMBER: US 60/088,241
/ PRIOR FILING DATE: 1998-06-05
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 2
/ LENGTH: 1542
/ TYPE: RNA
/ ORGANISM: Escherichia coli
US-09-465-355-2

```

Query Match	69.3%;	Score 995.2;	DB 3;	Length 1542;
Best Local Similarity	66.2%;	Pred. No. 0;		
Matches 969;	Conservative 255;	Mismatches 253;	Indels 7;	Gaps 5;

[illegible]

QY	127	GAAAGACCTTTCGCTTTTGGAGCGGCGCATGTCTGATTGACTAGTGTGGGGTTAAAGC	236
Db	208	UUCCGGCCCUCCUGGCACUCCGAGUGUGCCCAAGUGGGAUUAGUAGUAGUGUGGAGUAAACCGC	267
QY	237	CTACCAAGGCGCACATCGATGATGTGTGAGAGGACACACAGCCACACTGGGACTTGAGA	296
Db	268	UACACCUAGGCGACCAUCCCUAGCUUGUCUGAGAGAUACAAGCCACACUCUGAAUCUGAGA	327
QY	297	CACGGCCCACTCTCTACGGGAGGCGACGATGGGGAATTTTGGACATGGGCGCAAGCCT	356
Db	328	CACGGUCACAGCUCUACCGGAGGCGACGAGUGGGGAUUUUGSCAAUAGGCGCGCAAGCCU	387
QY	357	GATCCAGCAATGCCCGCTGAGTGAAGAAAGGCTTCGGGTTCTTAAGCTCTTTCAGTGGAG	416
Db	388	GAUCGACGCAAGCCGCGUGUAUAGAAAGGCTCUGGUGUUUAAAGUACUUACAGCGGGG	447
QY	417	AAGAAAGGTTACGCTTAAATATATCGTGACTCATGACGGTATTCGACAGAAAGACCGGCG	476
Db	448	AGGAAAGGAGUAAAGUUAUACUUUGCTUACUAGCUUACCCGCGAAGAAAGACCGGCG	507
QY	477	TAACTACGTGCGACAGCGCCGCGTAAATACGTAGGGGTGCAAGCGTTAATCGAAATTACTGG	536
Db	508	UAACTCCGUGCCAGACAGCGCGGUAUACGAGGGUGCAACGUUAAACGGAUUUACUGG	567
QY	537	CGCTTAAAGGGTCCGACGGCGGCTTTGTAATGCAATGTGAATTCGCCGGCTTAACTCGG	596
Db	568	GCUNAAAGCCGACCGACGGGUGUUAUAGUCAUUGGAAUCCCCGGGCTUCAAACUUG	627
QY	597	GAATTGGCTTGAATCTACAGGGCTAGAGTGTGGCAGAGGGAGGTGGAAATTCATGTGTGA	656
Db	628	GAUCUGAACUGUAUACUGGCAAGCUUGAGUCUCUGAAGGGGGGUAAGAUUCCAGUGUUA	687
QY	657	GCAGTGAATGCTTAGAGATATGGAAGAACATTCATGGCGGAAGCAGCGCTTCGGGTTAA	716
Db	688	GGGUGAAGUAGCGUAGAGAUUCUGAGAGAAUACCGGUGCGAAAGCGGCCCCUGGACGAA	747
QY	717	CACGTAGCGCTATGCAAGAAAGCGTGGGAGCAAAACAGATTAGTACCTCGTGTAGTCCA	776
Db	748	GACUGACGCUACAGUGCGAAAGCCUGGGGAGCAAAACAGAAUUAAUACCCUGGAGUUC	807
QY	777	CGCCCTTAAAGATGTCAACTAGTGTGGGGCTTATTAGAGCTTGGG-TAACAGAGCTAAGC	835
Db	808	CGCGGUAAACAUUGCCGACUUGAGAGUGUGCCCUUGAGGCGUGGCUUCCGAGCUAACG	867
QY	836	CGTAGAATGTGACCGCTCGGAGGATACCGTTCGCAAGATTTAAACTCAAAAGAAATGACGGG	895
Db	868	CGUUAAGUICGACCGCTGGGGAGUACCGCGCAAGUUAUAAACUCAAUUGAUUGACGGG	927
QY	896	GACCCGCGCAAGCGGTGATTAATGTGATTAATTCGATGCAACGCGGAAAAACCTTAACCTA	955
Db	928	GGCCCGCGCAAGCGGUGAGCAUUGUGUUUAUUCGAVGCAACCGGAAGAACCTUACCCUG	987
QY	956	CCCTTGAACATGTAGCGAATTTTCTTAGAGAT-AGATTAGTCTTGGGGGAACGCTAACAG	1014
Db	988	GUCUUGACAUCCACGGAAGUUUUUCHAGAGUAAGAAUUGGCUUCUGGGAACCGUAGACAG	1047
QY	1015	GTGCTGACATGCTGTCTGCACTCGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGC	1074
Db	1048	GUGUGUAGUGGCUUUCGUCAGCUCUGUGUUGUAGAAUUGUGGUUAUAGUCCCGCAACGAGC	1107
QY	1075	GCAACCCCTGTCTAATTAATGCGATC-ATTGCGTGGGCACTTTAATAGATCGCCGGTGA	1133
Db	1108	GCAACCCCUAACCUUUGUGUGCAGCGGCGCGCGGGAACUCAAAGGAGAGACUGGCCAGUGA	1166
QY	1134	CAAAACCGGAGAAAGTGGGGATGACGTCAAGATCCATAGCGCCMTTATGGGTAAGGGCTTCA	1193
Db	1168	UAAACUGGAGAAAGUGGGAGUAGCGUCAAUUCUACUGGCCCUUACGACCAAGGCUACA	1222
QY	1194	CACGTTAATACATAGTGGCGGTACAGAGGTTTCCAAACCGCGAGAGGGGAGCTAATCTCGA	1253
Db	1228	CACGUGUACAUAGGCCCGCAUACAAAGAAAGCAACUCGCGAAGAGCAAGCGGACUCUACA	1287
QY	1254	AAGGCGGTCTGATGTCGGATGCGAGTCTGCAACTCGACTCCGTGAAGTCCGAATCGCTAG	1313

DB 1288 AAGUCGCGUGUGUCGGAUUGGACUCGACUAGAAUGCGAUCGCG 1347  
 QY 1314 TAAATGCGGATCAGCATGTGCGGTGAATACGTTCCGGGTCTTGTACACACCGCCGTC 1373  
 DB 1348 UAAUGGUGAUGAUAUGCCACGUGAUAACUUCGCCGCGUUGAACACACCGCCGUC 1407  
 QY 1374 ACACCATGGAGTGGGTTTACACGAAGAGGTACTTACCGGTAAAGAGGCGCTTGC 1433  
 DB 1408 ACACCAUGGAGUGGUGUCAAAGAAUGAGUAGUAAACUUCGCGAGGCGCUAAC 1467  
 QY 1434 ACGGTGAGTTCATGACTGGGCTG 1457  
 DB 1468 ACUUGUGAUCUAGCUGGAGUG 1491

RESULT 37

US-08-520-946-158  
 ; Sequence 158, Application US/08520946  
 ; Patent No. 6372424  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BROW, MARY ANN D.  
 ; APPLICANT: LYMICHEV, VICTOR I.  
 ; APPLICANT: OLIVE, DAVID M.  
 ; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
 ; NUMBER OF SEQUENCES: 160  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSER: MEDLEN & CARROLL  
 ; STREET: 220 MONTGOMERY STREET, SUITE 2200  
 ; CITY: SAN FRANCISCO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: UNITED STATES OF AMERICA  
 ; ZIP: 94104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/520,946  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CARROLL, PETER G.  
 ; REGISTRATION NUMBER: 32,837  
 ; REFERENCE/DOCKET NUMBER: FORS-01756  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 705-8410  
 ; TELEFAX: (415) 397-8338  
 ; INFORMATION FOR SEQ ID NO: 158:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1542 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-08-520-946-158

Query Match 68.3%; Score 995.2; DB 3; Length 1542;  
 Best Local Similarity 82.2%; Pred. No. 0;  
 Matches 1204; Conservative 0; Mismatches 253; Indels 7; Gaps 5;

QY 1 ATTGAAGCTGCGGCGCATGCTTTACATGATGCAAGCGGACGACG--GATGCTTGC 58  
 DB 28 ATTGAAGCTGCGGCGGCGCTTAACATGATGCAAGCGTAAACGAAAGAGCTTGC 87  
 QY 59 ATC--TGGTGGCGAGTGGCGGACGGGTGATGATGATCGGAACGTATCCAGAGAGGG 116  
 DB 88 TTCCTTCTGACGAGTGGCGGACGGGTGATGATGCTGCGGAAACCTCGATGAGAGGG 147  
 QY 117 GGTAAAGCATGAAAGATGTCTTAATCCGATTAATCTTAAGAGGAAAGCGAGGATC 176

DB 148 GATACTACTGGAACCGTAGCTAATACCGCATTAACGTGCGAAGACCAAGAGGGGAGC 207  
 QY 177 GAAGACCTTCCGCTTTTGGAGCGCGCATGTCTGATTAGTATGTTGGGTAAAGGC 236  
 DB 208 TTCGGGCTCTTGGCATGATGTGCCCCAGATGGGATTTAGCTTAGTGGGTAAACGC 267  
 QY 237 CTACCAAGCGACATCAGTATGTTGCTGAGAGGACACAGCCACACTGSGACTGAGA 296  
 DB 268 TCACCTAGGCGACATCTCCTAGCTGTGTGAGAGGATGACACAGCCACACTGGAAC 327  
 QY 297 CACGGCCGACACTCTTACGAGGACGACGTGGGGAATTTTGGACAAATGGGCGCAAGCT 356  
 DB 328 CACGGTCCAGACTCTTACGAGGACGACGTGGGGAATTTTGGACAAATGGGCGCAAGCT 387  
 QY 357 GATCCAGCAATGCGCGGTGATGAAGAGGCTTCGGGTGTAAGCTCTTCAAGTCAG 416  
 DB 388 GATGACGACATGCGCGGTGATGAAGAGGCTTCGGGTGTAAGCTCTTCAAGTCAG 447  
 QY 417 AAGAAAGGTTACGGTAAATATGCTGATGATGATGATGATGATGATGATGATGATGAT 476  
 DB 448 AGAAGGAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 507  
 QY 477 TAACTAGTGCAGGACGCGGTAAATACGTAAGGTGCAAGCGTTAATCGAATTAATTC 536  
 DB 508 TAACTAGTGCAGGACGCGGTAAATACGTAAGGTGCAAGCGTTAATCGAATTAATTC 567  
 QY 537 GCGTAAAGGCGGACGCGGTAAATACGTAAGGTGCAAGCGTTAATCGAATTAATTC 596  
 DB 568 GCGTAAAGGCGGACGCGGTAAATACGTAAGGTGCAAGCGTTAATCGAATTAATTC 627  
 QY 597 GAAATGCGTTGAAACTCAAGGCTAGAGTGTGAGAGGAGGAGTGAATTCATGTGTA 656  
 DB 628 GAACTGATCTGATCTGAGCAAGCTTGTGATGATGATGATGATGATGATGATGATG 687  
 QY 657 GCAGTGAATGCGTGAAGATATGAAGAACATGATGATGATGATGATGATGATGATGAT 716  
 DB 688 GCGGTGAATGCGTGAAGATATGAAGAACATGATGATGATGATGATGATGATGATGAT 747  
 QY 717 CACTGACGCTCATGACAGAAAGCGTGGGAGGAAACAGGATTAATGATCCGTGTGTA 776  
 DB 748 GACTGACGCTCATGACAGAAAGCGTGGGAGGAAACAGGATTAATGATCCGTGTGTA 807  
 QY 777 GCGCCGTAAGCATGATCACTAGTGTGGGCTTATTAAGCTTGG--TAAAGAACTAACG 835  
 DB 808 CCGCGTAAAGCATGATGATCACTAGTGTGGGCTTATTAAGCTTGG--TAAAGAACTAACG 867  
 QY 836 CCGTGAAGTGAACGCGCTGGGAGTACCGTGGCAAGATTAACTCAAGAAATTGACGGG 895  
 DB 868 CCGTGAAGTGAACGCGCTGGGAGTACCGTGGCAAGATTAACTCAAGAAATTGACGGG 927  
 QY 896 GACCGGCAAGCGGTGATTAATGATGATTAATGATGATGATGATGATGATGATGATG 955  
 DB 928 GCGCGGCAAGCGGTGATTAATGATGATTAATGATGATGATGATGATGATGATGATG 987  
 QY 956 CCGTGAAGTGAACGCGCTGGGAGTACCGTGGCAAGATTAACTCAAGAAATTGACGGG 1014  
 DB 988 GCTTGAAGTGAACGCGCTGGGAGTACCGTGGCAAGATTAACTCAAGAAATTGACGGG 1047  
 QY 1015 GTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1074  
 DB 1048 GTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1107  
 QY 1075 GCAACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133  
 DB 1108 GCAACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167  
 QY 1134 CAACCGGAGAAAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1193  
 DB 1168 TAAACTGAGAGAAAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1227  
 QY 1194 CACGTAAATCAATGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253  
 DB 1228 CACGTAAATCAATGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1287



QY 1194 CACGTATATCATGCGGCTGTACAGAGGTTGCCAACCCGGGAGGCTAATCTCAGA 1253  
 DB 1228 CACGTGTACAAATGGCGCATACAAAGAGGAGCAGCTCGAGAGGACGAGCCTCATATA 1287  
 QY 1254 AAGCGGCTGTAGTCCGGATCGGAGTCTGCAACTCGACTCCGTGAAGTGGAAATCGCTAG 1313  
 DB 1288 AAGTGGTGTAGTCCGGATCGGAGTCTGCAACTCGACTCCGTGAAGTGGAAATCGCTAG 1347  
 QY 1314 TAATCGCGATCAGCATGTCTGGGTGAATACGTTCCCGGGCTTGTACACACCGCCGTC 1373  
 DB 1348 TAATGTGTATCAGAAATGCGAGGTGAATACGTTCCCGGGCTTGTACACACCGCCGTC 1407  
 QY 1374 ACACCATGGGAGTGGGTTTCCACAGAGCAGGTAAGTCTAACCGTGAAGAGGGCGCTTGGC 1433  
 DB 1408 ACACCATGGGAGTGGGTTTCCACAGAGCAGGTAAGTCTAACCGTGAAGAGGGCGCTTACC 1467  
 QY 1434 ACGGTGATTCATGACTGGGGTG 1457  
 DB 1468 ACTTGTGATTCATGACTGGGGTG 1491

RESULT 39

US-09-548-998B-33  
 Sequence 33, Application US/09548998B  
 Patent No. 6797817  
 GENERAL INFORMATION:  
 APPLICANT: E. I. DUPONT DE NEMOURS & COMPANY, INC.  
 APPLICANT: EBERSOLE, RICHARD C.  
 APPLICANT: HENDRICKSON, EDWIN  
 TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS FOR THE IDENTIFICATION OF DECHLORINATING E  
 FILE REFERENCE: BCI002 US NA  
 CURRENT APPLICATION NUMBER: US/09/548,998B  
 PRIOR FILING DATE: 2000-04-14  
 PRIOR APPLICATION NUMBER: 60/129,511  
 NUMBER OF SEQ ID NOS: 60  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 33  
 LENGTH: 1542  
 TYPE: DNA  
 ORGANISM: Escherichia coli  
 US-09-548-998B-33

Query Match 68.3%; Score 995.2; DB 3; Length 1542;  
 Best Local Similarity 82.2%; Pred. No. 0;  
 Matches 1204; Conservative 0; Mismatches 253; Indels 7; Gaps 5;

QY 1 ATGGAACGCTGGCGGATGCTTTACATGCAAGTGAACGCGACGCG--GATGCTTGC 58  
 DB 28 ATTGAACGCTGGCGGCGGCTTAACATGCAAGTGAACGCGTAACGGAAGAGCTTGC 87  
 QY 59 ATC--TGGTGGCGAGTGGCGGAGCGGTGAATTAATGCAATCGGAACGTATCCAGAAAGGGG 116  
 DB 88 TTTCTTGTCTGACGATGGCGGAGCGGTGAATTAATGCTGGAAATGCTCGATGAGGGG 147  
 QY 117 GGTAAAGCATGGAAGATGTCTTAATACCGCATATCTTAAGAGGGAAGACAGGGATC 176  
 DB 148 GATTAATCTAGGAACGCTTAATACCGCATATCTCGCAAGACCAAGAGAGGGGAGAC 207  
 QY 177 GAAAGACCTTGGCGTTTGGAGCGGCGGATGTCTGATTAAGTCTAGTGTGGGTAAAGGC 236  
 DB 208 TTGGGGGCTCTTGGCATCGAATGCGCAAGATGGGATTAAGTGAAGGCGTAAAGGC 267  
 QY 237 CTACCAAGGCGAGCATAGTAATGTTGTCTGAGAGGAGCAAGCAACTGGGAGCTGAGA 296  
 DB 268 TCACCTTGGCGGACATCTTGAATGTTGTCTGAGAGGAGCAAGCAACTGGGAGCTGAGA 327  
 QY 297 CACGCGCCAGACTCTTACGCGGAGGAGCAAGTGGGGAATTTTGAACAATGGGCGCAAGCT 356  
 DB 328 CACGCTCAAGACTCTTACGCGGAGGAGCAAGTGGGGAATTTTGAACAATGGGCGCAAGCT 387  
 QY 357 GATTCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGTAAAGCTCTTTCAGTGCAG 416

DB 388 GATGAGCCATGCGCGTGTATGAGAAGGCTTTCGGGTTGTAAAGTACTTTCAGCGGGG 447  
 QY 417 AAGAAAGGTTAAGGTTAATATATGCTGATGACGGTATCGACAGAAAGAGCAGCGGC 476  
 DB 448 AAGAAAGGAGTAATTAATTAATCTTGTGCTCATTTGACGTTACCCGCAAGAAAGACCGGC 507  
 QY 477 TAATACGTCAGAGCGCGGTTAATCTGTAAGGTGCAAGCGTTAATCGGAATTAATCTGG 536  
 DB 508 TAATCCGTGCGAGAGCGCGGTTAATTCGAGAGGTGCAAGGTTAATCGGAATTAATCTGG 567  
 QY 537 GGTAAAGGAGTGGCGAGCGGCTTGTGATGATGAGTGAATATCCCGGAGCTTAACTGG 596  
 DB 568 GGTAAAGGAGCGAGCGGCTTGTGATGATGAGTGAATATCCCGGAGCTTAACTGG 627  
 QY 597 GAATTCGCTTGAATCTAAGGCTTAAGTGTGGAGAGGAGGAGGAGGATTCATGATGTA 656  
 DB 628 GAATGCAATCTGATCTGCAAGCTTGGATGCTCGTAGAGGGGAGTAATTCAGAGTGA 687  
 QY 657 GCAGTGAATGCGGTGAGATTAAGAAAGCAATCGATGCGAAGGAGCGCTCTCGGATTAA 716  
 DB 688 GCGGTGAATGCGGTGAGATCTGAGAGAAATACCGGTGGAGAGGCGCCCTTGAGCGAA 747  
 QY 717 CACTGACGCTATGACAGAAAGCGTGGGAGCAACAGATTAATATCCTCTGTGTGCA 776  
 DB 748 GACTGACGCTATGAGTGGCAAGCGTGGGAGCAACAGATTAATATCCTCTGTGTGCA 807  
 QY 777 GCGCTTAAACGATGCTCACTAATGTTGGGCTTATTTAGGCTTGG--TAACGAAGCTTAAG 835  
 DB 808 CCGCGTAAACGATGCTCACTAATGTTGGGCTTATTTAGGCTTGG--TAACGAAGCTTAAG 867  
 QY 836 CGTGAAGTGAACGCGCTGGGAGTACGCGTGGCAAGTTAACTCAAGAAAGAAATTTGACGG 895  
 DB 868 CGTTAAGTGAACGCGCTGGGAGTACGCGTGGCAAGTTAACTCAAGAAAGAAATTTGACGG 927  
 QY 896 GACCGCAACAGCGGTGATTAATGATTAATTTGATGCAACGCGAAAGAACTTAAGCTTA 955  
 DB 928 GCGCGCAACAGCGGTGATTAATGATTAATTTGATGCAACGCGAAAGAACTTAAGCTTA 987  
 QY 956 CCGTTGACATGATGAGAAATTTCTTGAAGAT--AGATTAAGTCTTGGGAAACGCTAACAG 1014  
 DB 988 GTCTTGACATCAACGGAATTTTCAAGATGGAATGTGCTTGGGAAACGCTGAGACAG 1047  
 QY 1015 GTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1074  
 DB 1048 GTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1107  
 QY 1075 GCAACCTTGTCAATTAATGCAATC--ATTGCTTGGGCACTTTAAATGAGACTGCGGTGA 1133  
 DB 1108 GCAACCTTGTCAATTAATGCAATC--ATTGCTTGGGCACTTTAAATGAGACTGCGGTGA 1167  
 QY 1134 CAACCGAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193  
 DB 1168 TAACTGAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1227  
 QY 1194 CACGTATATCATATGAGGCGCTAAGAGGTGCGCAACCGCGAGGGGAGCTAATCTCAGA 1253  
 DB 1228 CACGTCTCAATATGAGGCGCTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1287  
 QY 1254 AAGCGGCTGTAGTCCGGATCGGAGTCTGCAACTCGACTCCGTGAAGTGGAAATCGCTAG 1313  
 DB 1288 AAGTGGTGTAGTCCGGATCGGAGTCTGCAACTCGACTCCGTGAAGTGGAAATCGCTAG 1347  
 QY 1314 TAATCGCATCAGCATGTCTGGGTGAATACGTTCCCGGGCTTGTACACACCGCCGTC 1373  
 DB 1348 TAATGTGTATCAGAAATGCGAGGTGAATACGTTCCCGGGCTTGTACACACCGCCGTC 1407  
 QY 1374 ACACCATGGGAGTGGGTTTCCACAGAGCAGGTAAGTCTAACCGTGAAGAGGGCGCTTGGC 1433  
 DB 1408 ACACCATGGGAGTGGGTTTCCACAGAGCAGGTAAGTCTAACCGTGAAGAGGGCGCTTACC 1467  
 QY 1434 ACGGTGATTCATGACTGGGGTG 1457  
 DB 1468 ACTTGTGATTCATGACTGGGGTG 1491

```

RESULT 40
US-10-061-071-33
; Sequence 33, Application US/10061071
; Patent No. 6894156
; GENERAL INFORMATION:
; APPLICANT: E. I. DUPONT DE NEMOURS & COMPANY
; APPLICANT: HENRICKSON, EDWIN
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS FOR THE IDENTIFICATION OF DECHLORINATING
; FILE REFERENCE: BCI1002 US CIP
; CURRENT APPLICATION NUMBER: US/10/061,071
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 60/129,511
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/129,511
; PRIOR FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-061-071-33

Query Match      68.3%; Score 995.2; DB 3; Length 1542;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1204; Conservative 0; Mismatches 253; Indels 7; Gaps 5;

QY 1 ATGTAACGCTGGCGGCGATCTTACACATGACGAAAGCGGACGAGC--GATGCTTGC 58
DB 28 ATGTAACGCTGGCGGCGGAGCTTAAACATGACGAAAGCGGTAACAGAGAGAGCTTGC 87
QY 59 ATC--TGGTGGCGAGTGGCGGAGCGGGTGAATGATCATCGAAACGTATCCAGAAAGGCG 116
DB 88 TTCTTTGGTGAAGAGTGGCGGAGCGGGTGAATGATGCTGGAAACCTGCTGATGAGGCG 147
QY 117 GGTAAACGATCGAAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGAGGGGATC 176
DB 148 GATTAATCTAGTGAACCGTAACTAACCGCATTAACGTGCAAGACCAAGAGGGGAGCC 207
QY 177 GAAACACCTGGCGCTTTGGAGCGGCGGATGCTGATTAAGTATGTTGGGGGTAAAGGC 236
DB 208 TTCCGGCCCTCTTGCATCGAGATGTCCTGATGAGATGTAAGTGAAGTGAAGCGC 267
QY 237 CTACCAAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 296
DB 268 TCACTTACGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 327
QY 297 CACGCGCCGAGATCTCTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 356
DB 328 CACGCTTCAAGATCTCTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 387
QY 357 GATTCAGCAATGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 416
DB 388 GATTCAGCAATGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 447
QY 417 AAGAAAGGTTACGCTTAATATCTGATCTCATGACGATGACGATGACGATGACGATGACG 476
DB 448 AGGAAGGAGGTTAATATCTTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 507
QY 477 TAACTACGTCAGACGCGCGGATTAATGATGATGATGATGATGATGATGATGATGATGAT 536
DB 508 TAACTACGTCAGACGCGCGGATTAATGATGATGATGATGATGATGATGATGATGATGAT 567
QY 537 GGTAAAGGAGTGGAGGCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 596
DB 568 GGTAAAGGAGTGGAGGCGGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 627
QY 597 GAATTCGCTTTGAATCAAGGCTTAAGATGATGATGATGATGATGATGATGATGATGATGAT 656
DB 628 GAATTCGCTTTGAATCAAGGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 687

```

```

QY 657 GCAGTAATGCGTGAAGATATGGAAGACATGATGCGGAGGAGGAGGAGGAGGAGGAGGAGG 716
DB 688 GCGGTAAATGCGTGAAGATATGGAAGATATGGAAGATATGGAAGATATGGAAGATATGGA 747
QY 717 CACTGACGCTTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 776
DB 748 GACTGACGCTTCATGACGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 807
QY 777 GCGCTTAAAGATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 835
DB 808 GCGCTTAAAGATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
QY 836 CGTGAAGTGAACGCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 895
DB 868 CGTGAAGTGAACGCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 927
QY 896 GACCCGACAAAGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 955
DB 928 GCGCCGACAAAGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 987
QY 956 CCTTGAATGATGACGAAATTTTCTAGAGAT-AGATTAGTGTGTTGGGAGGCTTAACAG 1014
DB 988 GTCTTGAATGATGACGAAATTTTCTAGAGAT-AGATTAGTGTGTTGGGAGGCTTAACAG 1047
QY 1015 GTGCTGATGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1074
DB 1048 GTGCTGATGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107
QY 1075 GCAACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133
DB 1108 GCAACCTTGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167
QY 1134 CAATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1193
DB 1168 TAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1227
QY 1194 CAGCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
DB 1228 CAGCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1287
QY 1254 AAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
DB 1288 AAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347
QY 1314 TAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1373
DB 1348 TAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1407
QY 1374 ACACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1433
DB 1408 ACACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1467
QY 1434 ACGGTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1457
DB 1468 ACTTGTGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1491

RESULT 41
US-09-726-774-2
; Sequence 2, Application US/09726774
; Patent No. 6677153
; GENERAL INFORMATION:
; APPLICANT: Iversen, Patrick L.
; TITLE OF INVENTION: Antisense Antibacterial Method and
; FILE REFERENCE: 0450-0032.30
; CURRENT APPLICATION NUMBER: US/09/726,774
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 60/166,150
; PRIOR FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 2  
; LENGTH: 1541  
; TYPE: DNA  
; ORGANISM: Salmonella thymimurium  
US-09-726-774-2

Query Match 68.1%; Score 991.8; DB 3; Length 1541;  
Best Local Similarity 81.7%; Pred. No. 0;  
Matches 1195; Conservative 0; Mismatches 262; Indels 6; Gaps 4;

```
QY 1 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTGAACGG---CAGACGGATGCTTG 57
DB 28 ATTGAACGCTGGCGGCGGCGATCTTTACATGCAAGTGAACGGTAAACGAGACGCTTTC 87
QY 58 CATCTGGTGGCGAGTGGCGGAGCGGCTGAATGCAATCGGAAAGTATCCAGAGAGAGGGG 117
DB 88 TCTTTGCTGACGAGTGGCGGAGCGGCTGAATGCTGGGAAACTGCCATGAGAGGGG 147
QY 118 GTAAAGCATGAAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGAGTGC 177
DB 148 ATTAATCTAGGAAACGGTGGCTTAATACCGCATATATCTTAAGAGAGAAAGCAGGGAGTGC 207
QY 178 AAAAGACTTGGCGCTTTTGAAGCGGCGGATGTCTGATTAAGTGAAGTGGGTTAAAGGCC 237
DB 208 TCGGGCTCTTGGCATGCGATGTGCGCAATGGGATTAAGTGAAGTGGGTTAAAGGCC 267
QY 238 TACCAAGGAGGAGATGATGATGCTGAGAGAGAGCAGCAGCAGCTGAGAGCTGAGAC 297
DB 268 CACTTAAGGAGAGATCTCTTAAGTGTGCTGAGAGAGAGCAGCAGCAGCTGAGAGCTGAGAC 327
QY 298 ACGGCGCAGACTCTTAAGGAGAGCAGCAGTGGGAAATTTGCAATAGGCGGAGCCTTG 357
DB 328 ACGGTCTCAAGCTCTTAAGGAGAGCAGCAGTGGGAAATTTGCAATAGGCGGAGCCTTG 387
QY 358 ATCCAGCAATGCCGCTGAGTGAAGAGCGCTTCCGCTTGAAGCTCTTTCAGTGAAGA 417
DB 388 ATGCAAGCAGTGCCTGTATGTAAGAGAGCGCTTCCGCTTGAAGCTCTTTCAGTGAAGA 447
QY 418 AGAAAGGTTAAGGTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 477
DB 448 GGAAGGTTGTTGTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
QY 478 AACTAGTGCACAGCAGCAGCGGATTAATGATGATGATGATGATGATGATGATGATGATGAT 537
DB 508 AACTCGGTGCACAGCAGCAGCGGATTAATGATGATGATGATGATGATGATGATGATGATGAT 567
QY 538 CGTAAAGGTTGCGCAGGCGGCTTTGTAAGTCAATGTTGAATCCCGGCTTAACTGCGG 597
DB 568 CGTAAAGGTTGCGCAGGCGGCTTTGTAAGTCAATGTTGAATCCCGGCTTAACTGCGG 627
QY 598 AATTTGGTTTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
DB 628 AACTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
QY 658 CAGTGAATATCGTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
DB 688 CGGTGAATATCGTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
QY 718 ACTGACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAAGTATCCCTGATGATGAT 777
DB 748 ACTGACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAAGTATCCCTGATGATGAT 807
QY 778 GGCCTTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 836
DB 808 GCGGTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
QY 837 GTGAAGTTGACCGGCTGGGAGTACGGTCCGAGATTAATCAAAAGATTAAGTGAAGGAG 896
DB 868 GTTAAGTGAAGTGTGGGAGTACCGGCTGGGAGTAAATCAAAAGATTAAGTGAAGGAG 927
QY 897 ACCGCAAGAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAATCTTAATCTTAC 956
DB 928 GCCGCAAGAGCGGTGATGATGATGATTAATTCGATGCAACGCGAAATCTTAATCTTAC 987
```

```
QY 957 CTTGACATGATGAGGAATTTTCTAGAGAT--AGATTAGTCTTCGGGAAACGCTAACAGAG 1015
DB 988 TCTTGACATGACAGAACTTTTCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1047
QY 1016 TGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1075
DB 1048 TGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107
QY 1076 CAACCTTGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134
DB 1108 CAACCTTATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167
QY 1135 AAACCGGAGAAAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1194
DB 1168 AAACCGGAGAAAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1227
QY 1195 AGCTAATCAATGAGGCGGATGAGAGGTTGCAACCGCGAGAGGAGGAGCTAATCTCAGAA 1254
DB 1228 AGCTCTCAATGAGGCGGATGAGAGGTTGCAACCGCGAGAGGAGGAGCTAATCTCAGAA 1287
QY 1255 AGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
DB 1288 AGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347
QY 1315 AATCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1374
DB 1348 AATCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1407
QY 1375 CACCATGAGAGGAGGTTTCAACAGAGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1434
DB 1408 CACCATGAGAGGAGGTTTCAACAGAGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1467
QY 1435 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1490
DB 1468 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
```

RESULT 42  
US-09-726-774-4  
; Sequence 4, Application US/09726774  
; Patent No. 6577153  
; GENERAL INFORMATION:  
; APPLICANT: Iversen, Patrick L.  
; TITLE OF INVENTION: Antisense Antibacterial Method and  
; TITLE OF INVENTION: Composition  
; FILE REFERENCE: 0450-0032.30  
; CURRENT APPLICATION NUMBER: US/09/726,774  
; PRIOR FILING DATE: 2000-11-29  
; PRIOR APPLICATION NUMBER: US 60/168,150  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 1500  
; TYPE: DNA  
; ORGANISM: Vibrio cholera  
US-09-726-774-4

Query Match 67.9%; Score 989.4; DB 3; Length 1500;  
Best Local Similarity 81.8%; Pred. No. 0;  
Matches 1200; Conservative 3; Mismatches 254; Indels 10; Gaps 5;

```
QY 1 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTGAACGGAGCA---CGATGCTTG 57
DB 1 ATTGAACGCTGGCGGCGGCGATCTTTACATGCAAGTGAACGGAGCAATTTCAAAAGCTTG 60
QY 58 CATCTG---GTGCGAGTGGCGGAGCGGCTGAATGCAATCGGAAGTATCCAGAGAGG 114
DB 61 CTTTGAAGATGACGAGCGGCGGAGCGGCTGAATGCGCTGGAGACCTCCCTGAGCTGG 120
QY 115 GGGGTAAAGCATGAAAGATGCTTAATCCGATTAATCTTAAGAGGAAAGCAGGGGA 174
```



Db 121 GGGATTAAGGTTGAAAAGCACTGCTAATACCGCATGATGTTTACGAGCAACAAAGAGGGGGA 180  
Qy 175 TCGAAAAGACTTGGCGCTTTTGGAGCGGCCGATGTCTGATTAAGTGGTGGGTAAG 234  
Db 181 TTTTGGAGCTTTCGCTCGGGATGGGCCAGTTGGGATTACCTGATGGTGAAGTAATG 240  
Qy 235 GCTTACCAAGGGAGGATCAGTATGTTGGCTGAGAGGAGCAACAGCACTGGGACTGA 294  
Db 241 GTTCAACCAAGGGAGGATCCTTACCTGATGTTTGAAGATGATCAGCCACACTGGAATGA 300  
Qy 295 GACACGGCCAGACTCTTAACGGAGGAGCAAGTGGGGAATTTTGAACAATGGGCGCAAGC 354  
Db 301 GACACGGCTCAGACTCTTAACGGAGGAGCAAGTGGGGAATTTTGAACAATGGGCGCAAGC 360  
Qy 355 CTGATCCAGCAATGCCGCTGATGTAAGAAAGCCCTTCGGGTTTGAAGCTCTTTCAATCG 414  
Db 361 CTGATCCAGCAATGCCGCTGATGTAAGAAAGCCCTTCGGGTTTGAAGCACTTTTCAGAG 420  
Qy 415 AGAAGAAAAGGTTAGGTAATATCGTACTCATGACGGGTATCGACAGAAAGACACCG 474  
Db 421 TGAGAGAGGTTGGTCGTTAATAGCGTATCAATTGACCTTAGCTGCAGAGAGACACCG 480  
Qy 475 GCTAATCACTGTCAGACGCCCGGTAATACGTAGGGTGCAGCGTTAATCGGAATTAAT 534  
Db 481 GCTAATCTCGTGCAGACGCCCGGTAATAGAGAGGGTGCAGCGTTAATCGGAATTAAT 540  
Qy 535 GGGCGTAAAGGGTGGCGAGCGGCTTTTGAATCAAGTGTGAATCCCGGGCTTAACCT 594  
Db 541 GGGCGTAAAGCGCATGACAGCGGTTTGTTAACCAAGATGTGAACCCCGGGCTCAACCT 600  
Qy 595 GGGAAATTCGTTTGAACCTAACAAGGCTAGTGTGGCAGAGGGAGTGGGAATTCATG 654  
Db 601 GGGAAACCGCATTTTGAACCTGCAAGGCTAGAGCTTGTAGAGGGAGTGGGAATTCAGTG 660  
Qy 655 TAGCAGTAAATGCGTGAATGATGGAAGAACATGATGCGAGGACGCTCTGGGTT 714  
Db 661 TAGCGTGAATGCGTGAATGATGGAAGAACATGATGCGAGGACGCTCTGGGTT 720  
Qy 715 AACACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTGATCCTTGGTATGTC 774  
Db 721 AAGACTGACGCTCATGATGCAAGAGCGTGGGAGCAACAGATTGATCCTTGGTATGTC 780  
Qy 775 CACGCCCTAAACGATGCAAC-TAGTGTGTGGGCTTATTAGGCTTGGTAACGAAAGCTTA 833  
Db 781 CACGCTGAACGATGCTTACTTGAAGGTTGACCTTATAGTCTGTGGCTTTGGGCTPA 840  
Qy 834 CGCGTGAAGTTGACCGGCTGGGAGTACGCTGCAAGATTAAACTCAAGGAATTGACG 893  
Db 841 CGCGTGAAGTTGACCGGCTGGGAGTACGCTGCAAGATTAAACTCAAGGAATTGACG 900  
Qy 894 GGGACCCGACAAAGCGGTGATTTATGTGATTAAATTCAGTACGCGAAGAAACCTTAAC 953  
Db 901 GGGGCCCCACAAAGCGGTGAGCATGTGTTAAATTCATGACGCGAAGAAACCTTAAC 960  
Qy 954 TACCTTGAACATGTAGCGAATTTTCTAGAGATGATTTAGT-CTTCGGGAAACGCTAACAC 1012  
Db 961 TACTTTGAACATCCAGAGAAAGCGAAGAGATTTGTGTGCTTCGGGAAATCTGAAGAC 1020  
Qy 1013 AGGTCTCAGTGGCTGTGCTGAGCTGCTGTGAGATGTTGGTAAAGTCCCGCAAGCA 1072  
Db 1021 AGGTCTCAGTGGCTGTGCTGAGCTGCTGTGAGATGTTGGTAAAGTCCCGCAAGCA 1080  
Qy 1073 GCGCAACCCCTGTCTTAATTTGCAATC--ATTGGTTGGGCACTTTAAATGAGCTGCCG 1130  
Db 1081 GCGCAACCCCTTAATCTTGTGTTGCCAGCGATATTCGGGAACTCAAGAGAGACTGCCG 1140  
Qy 1131 TGAACAACCGAGAGAGGTGGGAGTGAAGTCAAGTCTCAAGGCCCTTAATGGGTAGGCT 1190  
Db 1141 TGAATAACCGAGAGAGGTGGGAGTGAAGTCAAGTCAATGAGCCCTTAATGAGTGGGCT 1200  
Qy 1191 TCACAGTAATACATGAGCGGCTGACAGAGGTTGCAACCGCGAGGGGAGACTAATCTC 1250  
Db 1201 ACACAGCTGCTAATGAGCATATACAGAGGCGAGGAGCGCGAGGTGAGAGCAATCCC 1260

Qy 1251 AGAAGCGCGTGTAGTCCCGGATCGAGTCTGCAACTCGACTCCGTAAGTCCGAAATCCG 1310  
Db 1261 AGAAGTATGTGTAGTCCGATCGAGTCTGCAACTCGACTCCGTAAGTCCGAAATCCG 1320  
Qy 1311 TAGTAATCGCGGATCAGATGCTCGCGGTGAATTAAGTCCGCGGCTTTGTACACACCGCC 1370  
Db 1321 TAGTAATCGGATCAGATGCTCAGGTAATGCTCGCGGCTTTGTACACACCGCC 1380  
Qy 1371 GTCAACACCGAGTGGGTTTCAACGAGACAGTGTCTAACCTGAAGAGGCGCTT 1430  
Db 1381 GTCAACACCGAGTGGGTTTCAACGAGACAGTGTCTAACCTGAAGAGGCGCTT 1440  
Qy 1431 GCCACGGGAGATTCAATGACTGGGCTG 1457  
Db 1441 ACCACGGTGTTCATGACTGGGCTG 1467

RESULT 43  
US-09-737-297-1  
; Sequence 1, Application US/09737297  
; Patent No. 6887984  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Mark  
; APPLICANT: Griffiths, Allen  
; APPLICANT: Hill, Philip  
; APPLICANT: Laybourne-Parry, Johanna  
; APPLICANT: Mills, Sarah  
; TITLE OF INVENTION: Processes and Organisms for the Production of Antifreeze Proteins  
; FILE REFERENCE: F3247  
; CURRENT APPLICATION NUMBER: US/09/737,297  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: GB 9929696.4  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 1486  
; TYPE: DNA  
; ORGANISM: Marlinomonas protea  
US-09-737-297-1

Query March 67.9%; Score 989.2; DB 3; Length 1486;  
Best Local Similarity 82.4%; Pred. No. 0;  
Matches 1207; Conservative 0; Mismatches 248; Indels 9; Gaps 6;

Qy 1 ATTGAAGCTGGCGGATGCTT-TACACATGCAAGTGAACCGGACGACGAGTCTTGCA 59  
Db 11 ATTGAAGCTGGCGGATGCTTAAACATGCAAGTGAAGCGGTGAACAGGGGCTTGT 70  
Qy 60 TCTGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGGAACGTATCCAGAAAGGGGCT 119  
Db 71 CTTCTGACGAGCGCGGACGGGTGAGTAATGCGCTAAGAAATCTGCTAGTAGAGGGGAC 130  
Qy 120 AAGCATCGAAAGATGTGCTAATACCGCATTAATCTTAAGAGGAAAGCAGGGATGAA 179  
Db 131 AAGATGTGAAACCGCATGCTTAATACCGCATACCTTGAAGGAGAAAGAGGGGACTTT 190  
Qy 180 AG--ACCTGGCTTTTGAAGCGGCGCATGTCTGATTAAGTATGTTGGTGGGTAAGGCC 237  
Db 191 CGAGGCTTCGGCTATTAGATGAGCCCTGCGTGAATTAAGTATGTTGGTAAAGGCC 250  
Qy 238 TACCAAGGCGAGTCAATGATGTTGTTGAGAGGACCAACGACCACTGGGACTGAGAC 297  
Db 251 TACCAAGGCGAGTCTTAATCTGATGTTGAGAGGATCAACGATCACTGGGACTGAGAC 310  
Qy 298 ACCGCGCAAGCTCTTAAGGGGAGGAGAGTGGGGAATTTTGAACAATGGGCGCAAGCTG 357  
Db 311 ACCGCGCAAGCTCTTAAGGGGAGGAGAGTGGGGAATTTTGAACAATGGGCGCAAGCTG 370  
Qy 358 ATCCAGCAATGCGCGTGAAGTGAAGAGGCTTCGGGTTGTAAGTCTTTCACTGAGAGA 417  
Db 371 ATCCAGCAATGCGCGTGAAGTGAAGAGGCTTTAGGTTTGAAGCACTTTCAAGGGTGA 430

418 AGAAAAAGTTACGTAATATATGTCATGACGGTATCGACAGAAAGACCGGCT 477  
431 GGAAGGGTATAGTTAATACGTATCATCTGACGTTAGCCCAAGAAAGACCGGCT 490  
478 AACTACGTCCACAGACCGCGGTAAATAGTATGGGTGCAAGCGTTATCGGAATTACTGG 537  
491 AACTGTGTCCACAGACCGCGGTAAATAGAGGGGTCCAGCGTTATCGGAATTACTGG 550  
538 CGTAAAGGGTGGCAGACCGCGTTGTAAGTCAGATGTAATCCCGGGCTTAACCTGG 597  
551 CGTAAAGCGCGGTAGGTGTTGTTAAGTCGATGTGAATCCAGGGCTCAACCTGG 610  
598 AATGCGTTGAACTCAAGAGCTAGAGTGTGGCAGAGGGAGGTGAATTCATGTGAG 657  
611 AATGGACCCGATACCTGAGTATGATGTAAGAGGGGTGTGAATTTCTGTGAG 670  
658 CAGTGAATGCGTATAGATATGGAAGAACATCGATGGCGAAGGCAAGCTCTGGTTAC 717  
671 CGGTAAATGCGTATAGATATGGAAGAACATCGATGGCGAAGGCAAGCTCTGGTTAC 730  
718 ACTGACGCTCATGACAGAAAGCGTGGGGAGCAAGATATGATACCTGAGTATCCAC 777  
731 ACTGACGCTCATGAGTGGCAAGCGTGGGGAGCAAGATATGATACCTGAGTATCCAC 790  
778 GCGCTTAAAGCATGCTCACTAGTGTGGGCTTATTAAGGCTTGTGAAGCAAGCTAAC 837  
791 GCGCTTAAAGCATGCTCACTAGTGTGGGCTTATTAAGGCTTGTGAAGCAAGCTAAC 849  
838 TGAAGTTGACCGCTGGGGAGTACCGTCCGAAATTTAACTCAAGAAATTTAGCGGGA 897  
850 ATATAGTAGACCGCTGGGGAGTACCGTCCGAAATTTAACTCAAGAAATTTAGCGGGA 909  
898 CCGGCAAGAGCGGTGATATATGGAATTTATGATGCAAGCGGAAACCTTAACCTAC 957  
910 CCGGCAAGAGCGGTGATATATGGAATTTATGATGCAAGCGGAAACCTTAACCTAC 969  
958 CTGACATGATGCAAGATTTTCTAGAGT-AGATTAGTCTTCCGGAACGCTTAACAGAGT 1016  
970 CTGACATGATGCAAGATTTTCTAGAGT-AGATTAGTCTTCCGGAACGCTTAACAGAGT 1029  
1017 GCTGATGCTGTGCTGACGCTGTGCTGATGATGTTGGTTAAGTCCGCAACAGAGCG 1076  
1030 GCTGATGCTGTGCTGACGCTGTGCTGATGATGTTGGTTAAGTCCGCAACAGAGCG 1089  
1077 AACCTTGTGATTAATTTGCAATCA--TTTGGTTGGGCACTTTATGATGATGCTGCGGTAG 1134  
1090 AACCTTGTGCTTAATTTGCAATCA--TTTGGTTGGGCACTTTATGATGATGCTGCGGTAG 1149  
1135 AAACCGGAGAGAGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGAGGTAGGGCTTAC 1194  
1150 AAACCGGAGAGAGTGGGAGTGAAGTCAAGTCTCAATGCGCTTATGAGGTAGGGCTTAC 1209  
1195 ACCTAATATCAATGCGCGGTACAGAGGGTCCAAACCGGCAAGGGGAGCTTAATCTCAGAA 1254  
1210 ACCTGCTAATGCGCGGTACAGAGGGTCCAAACCGTACGATGAGTGAAGTCCCAACAA 1269  
1255 AGCGGCTGATGCTGCGGATGCGAGTCTGCAACTGCACTCGTGAAGTGGGAATCGCTAGT 1314  
1270 AGTACGCTGATGCTGCGGATGCGAGTCTGCAACTGCACTCGTGAAGTGGGAATCGCTAGT 1329  
1315 AATGCGGATCAGCATGCGCGGTGAATTAAGTCTCCGGGTCTTGTATACACACCGGCTTAC 1374  
1330 AATGCTGAATCAAGATGTCACGCTGAATGCTTCCGGGCTTGTATACACACCGGCTTAC 1389  
1375 CACATGAGAGTGGGTTTCAACAGAGCGAGTATCTAACCGT--AAGAGGGGCGCTTGC 1432  
1390 CACATGAGAGTGGGTTTCAACAGAGCGAGTATCTAACCGT--AAGAGGGGCGCTTGC 1449  
1433 CACGCTGATGATGATGCTGGGCT 1456  
1450 CACGAGTGTGATGATGCTGGGCT 1473

RESULT 44  
US-08-114-695A-1  
; Sequence 1, Application US/08114695A  
; Patent No. 5508193  
; GENERAL INFORMATION:  
; APPLICANT: Mandelbaum, Raphael T.  
; TITLE OF INVENTION: DEGRADATION OF S-TRIAZINES IN SOIL AND  
; TITLE OF INVENTION: WATER  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; STREET: 3500 IDS CENTER  
; CITY: MINNEAPOLIS  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/114,695A  
; FILING DATE: 31-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MUEJING, ANN M.  
; REGISTRATION NUMBER: 33,977  
; REFERENCE/DOCKET NUMBER: 600.268U51  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-339-0331  
; TELEFAX: 612-339-3061  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1542 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: rRNA  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli  
; US-08-114-695A-1  
Query Match 67.7%; Score 985.8; DB 2; Length 1542;  
Best Local Similarity 66.0%; Pred. No. 0;  
Matches 967; Conservative 232; Mismatches 257; Indels 9; Gaps 5;  
1 ATTGAACGCTGGCGGATCTTTACATGACATGCAAGTCAAGCGGCAAGCGGATGCTTGCAT 60  
28 AATGAACGCTGGCGGATCTTTACATGACATGCAAGTCAAGCGGCAAGCGGATGCTTGCAT 87  
61 C-----TGGTGGCGAGTGGCGGAGCGGGTGAATATGATCGGAACGATATCCAGAAAGGG 115  
88 CUUCUUGGCUAGAGAGUUGCGGAGCGGUGAGUUAUUCUGGGAACUUGCTCUAGAGGG 147  
116 GGGTAACGATCGGAAGATGCTAATATACCGATATATCTTAAGAGAGAAAGCAGGGGAT 175  
148 GGAUAAACUACUGGAAGACGUAUACGGAUAAACGUCGCA-GACCAAAAGAGGGGAGC 206  
176 CGAAAGACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAAGTATGTTGGGTAAAG 235  
207 CUUCGGGCUUCUUGCCAUUGGCUAGUUGCCGAGUUGAUCUAGUAGGUGGCUAAGCG 266  
236 CTTACCAAGGCGAGATGATGATGTTGGTGAAGAGGACGACCAAGCAGCACTGGGACTGAG 295  
267 CUACCUAAGCGAGUUGCCUAGUUGGCUAGUUGGCUAGUUGGCUAGUUGGCUAAGCG 326  
296 ACAAGGCGGAGACTCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 355  
327 ACAAGGCGGAGACTCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 386

```
QY 356 TGATCCAGCAATGCCGCTGAGTGAAGAAAGCCTTCGGGTGTGAAGCTCTTTCAGTCCA 415
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 387 UGAUGAGCCAUCCCGCGUGUAUGAAGAGCCUUCGGUUGUAAGUAUCUUACAGCGGG 446
QY 416 GAAAGAAAGGTTACGGTAAATATCTGTGACTCATGACGGTATGCAGAGAAAGCACCGG 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 447 GAGGAAAGGAGUAAGUAUAUACUUUGCUCAUUGACUUACCCGAGAAAGCACCGG 506
QY 476 CTAACTACGTCGACGAGCGCGGCTAAATACGTAGGGGTGCAAGGCTTAACGGAATTACG 535
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 507 CUAACUCCCGGACGAGCGCGGUAUUCGAGGGGUGCAAGCGUUAUCGGAUUUAUCUG 566
QY 536 GCGCTAAAGGGTGCAGCGCGGCTTGTAGTCAAGTGTGAATCCCGGCGCTTAACCTG 595
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 567 GCGGUAAACCGCACGAGCGCGUUCUUAAGUAGUUGAUAUCCCGGGGCUCAACUG 626
QY 596 GGAATTGCTTTGAACTCAAGGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGT 655
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 627 GGAACUGCAUCUGAUCUGGCAAGCUUGAGUCUGUAAGGGGGGUAUAUUCACAGUGU 686
QY 656 AACAGTGAATTCGCTAGATATGAGAAACATCGATGCGGAGGAGCGCTCTGGCTTA 715
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 687 ACCGGGAUUCGUGAGAUUCUGAGAGAAUACCGGUGCGGAGCGCCCTUGGACGA 746
QY 716 ACACGTACGCTATGCAGGAAAGCGTGGGAGCAAAACAGATTATGATACCTGTATGCTC 775
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 747 AACACUACCCUACGUGUGCCAAAGCGUGGAGCAAAACAGAUUAUACCCUGUAUCC 806
QY 776 ACCGCTTAACGATGTCAACTGATTGTTGGGCTTATTAAGCTTGG-TAACGAGCTAAC 834
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 807 ACCGCGUAAACGAUUGCGACUUGGAGUUGGCCUUGAGGCGUGGCUCCGAGCUAAC 866
QY 835 GCGTGAAGCTTGAACCGCTGGGAGTACCGGTCCCAAGATTAAATCAAGAAATGACGG 894
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 867 GCGUUAAGUGCAACCGCTUGGAGUAUCGCGCGCAAGUUAACUAAUUAUUGACGG 926
QY 895 GGAACCGCACAGCGGTGATTAATGATTAATTGATGCAACGCGAAACCTTACCT 954
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 927 GGGCCCGCACAGCGGUGAGCAUUGUGUUUAUUCGAGCAACGCGAAAGACCUUACCU 986
QY 955 ACCCTTGACATGTAGCGAAATTTTCTAGAGAT-AGATTAGTCTTGGGAAACGCTAACACA 1013
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 987 GGUUCUUGAACUCCAGAGUUAUUUCAGAGUAGAUUGGCUUCGCGAACGUAAGACA 1046
QY 1014 GGTGCTGCATGCTGCTGCTCACTGCTGCTGCTGATGTTGGGTTAAAGTCCCGCAACGAG 1073
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1047 GGUUGUGCAUUGGCUUGCAUCUCUGUUGAAUUGUGGUAUAAGCCCGCAACGAG 1106
QY 1074 CGCAACCTTGTCAATTAATGTCATC-ATTGATTGGGCACTTATATGAGATGCGCGTG 1132
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1107 CGCAACCTCUUACCUUUGUUGCCACGCGUCCGCGCGGAACTCAAGAGAGACUCCAGUG 1166
QY 1133 ACAAAACCGAGAGAGGTGGGATGACGTCAAGTCTCATGAGCCCTTATGGATGAGGCTTC 1192
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1167 AUAAAACUGGAGGAAAGUUGGAGUAGUACUACUACUACUACUACUACUACUACUACUAC 1226
QY 1193 ACACTAATACAAATGCGCGCTACAGAGGTTTCCAAACCGCGAGAGGAGACTAATCTCAG 1252
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1227 ACACUUGCUACAAUUGGCGCAUACAAAGAACGCAACUUGCGAGAGCAAGCGCAACUUCU 1286
QY 1253 AAAGCGCGTCTAGTCCGGAATCGAGTGTGCAACTCGACTCGGTGAGAGTCCGATCGCTA 1312
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1287 AAAGUGCGUUGUAGUCCGGAUUGAGUUCUGCAACUCCAGUAGUAGUAGUAGUAGUAGUAG 1346
QY 1313 GTAATCGCGATCAGCATGTCCGCGTGAATACGTTCCCGGATCTTGTACACACCGCCGT 1372
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1347 GUAUUGUGUAGUACAGAUAGCAACGUGUAGUAUACGUUCCCGGCGCUUGUACACACCGCCGU 1406
QY 1373 CACACCAATGGAGTGGGTTTCACAGAGACAGTGTCTAACCTGTAAGGAGGCGCTTGC 1432
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1407 CACACCAUUGGAGUGUGUUGCAAAAGAGUAGUAGCUUACCUCCGAGAGGCGCUUAC 1466
QY 1433 CACGGTGAAGATTCAATGACTGGGGTG 1457
```

```
Db 1467 CACUUGUAGUACUAGACUUGGUGU 1491

RESULT 45
US-09-492-709A-89
; Sequence 89, Application US/09492709A
; Patent No. 6720139
; GENERAL INFORMATION:
; APPLICANT: ZykInd, Judith
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; TITLE OF INVENTION: ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001A
; CURRENT APPLICATION NUMBER: US/09/492,709A
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 89
; LENGTH: 1549
; TYPE: DNA
; ORGANISM: E. coli
US-09-492-709A-89

Query Match 67.5%; Score 984.2; DB 3; Length 1549;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1204; Conservative 0; Mismatches 253; Indels 8; Gaps 6;

QY 1 ATGAAAGCTGGCGCGCATGCTTTAACAATGCAAGTCAACGCGACACGAT--GCTTGC 58
Db 28 ATGAAAGCTGGCGCGCATGCTTTAACAATGCAAGTCAACGCGATGCAAGGAGCTTGC 87
QY 59 ATCT--GGTGGCGATGGCGGAGCGGTGAGTAATGCAATCGGAATGCAAGAGAGGGG 116
Db 88 TGGCTCGTGAAGAGTGGCGGAGCGGTGAGTAATGCTGGGAACCTGCTGATGAGAGGG 147
QY 117 GGTAAAGCATGAAAGATGCTAATACCGCATATTAATCTTAAGAGAGAAAGCAGGGATC 176
Db 148 GATTAATCTAGGAACGTAAGTACTAATACCGCATATGCTGGAACCAAGAGGGGAGCC 207
QY 177 GAAAGACTTGGCGCTTTTGAAGCGCGGATGCTGATTAGCTAATGTTGGTGGGTTAAAGC 236
Db 208 TTGGGCGCTTGGCATTCGATGATGCGGATGCGGATGCTTGTGGGTTAAAGCGG 267
QY 237 CTACCAAGCGGACGATCAATGATGCTGATGAGAGAGACGACGACACTGGGACTGAGA 296
Db 268 TCACCAAGCGGACGATCCTAGTGTGCTGAGAGAGTACGACGACACTGGAACTGAGA 327
QY 297 CACGGCCAGACTCTCAACGAGGAGCAGACAGTGGGAAATTTTGAACAATGGGCGCAACCT 356
Db 328 CACGGTCCAGACTCTCAACGAGGAGCAGACAGTGGGAAATTTTGAACAATGGGCGCAACCT 387
QY 357 GATCCAGCATGCGCGGTGATGAGAAAGGCTTTCGGGTTGAAAGCTCTTTCAGTGCAG 416
Db 388 GATCCAGCATGCGCGGTGATGAGAAAGGCTTTCGGGTTGAAAGCTCTTTCAGTGCAG 447
QY 417 AAGAAAGGTTACGGTAAATATGTAATCTGATCATGACGCTATCCAGACAAAGAACACCGGC 476
Db 448 AAGAAAGGATGAAATTAATCTTGTCTCATTTGACCTTACCCGCAAAAGAACACCGGC 507
QY 477 TAACTACGTGACAGACGCGGCTGATATACGAGGCTGCAAGCTTATTCGAATTAATCTGG 536
Db 508 TAACTCCGTGACAGACGCGGCTGATATACGAGGCTGCAAGCTTATTCGAATTAATCTGG 567
QY 537 GCGTAAAGGTCGACAGCGGCTTTC--TAACTCAGATGTGAATCCCGGCGCTTAACCTG 595
Db 568 GCGTAAAGCGACAGCGGCTTTC--TAACTCAGATGTGAATCCCGGCGCTTAACCTG 627
```

596 GGAATTCGCTTTGAAACTCAAGGCTAGAGTGTGCAAGGAGGAGGATTCCTGCTGT 655  
628 GGAATTCGCTTTGAAACTCAAGGCTAGAGTGTGCAAGGAGGAGGATTCCTGCTGT 687  
656 AGCACTGAAATGCGTGAAGATTTGGAAGAACTGATGGCGAAGGAGGCTGCGGTTA 715  
688 AGCGGTGAATGCGTGAAGATTTGGAAGAACTGATGGCGAAGGAGGCTGCGGTTA 747  
716 ACATGAGCGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTAGTCC 775  
748 AGACTGAGCGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTAGTCC 807  
776 AGCGCTTAAACGATGCACTGATTTGTTGGGCTTATTAGGCTTGG-TAACGAAGCTAAC 834  
808 AGCGGTGAATGCGTGAAGATTTGGAAGAACTGATGGCGAAGGAGGCTGCGGTTA 867  
835 GCGTGAAGTTGACCGGCTGCGGAGTACCGGTGCGAAGATTAAACTCAAGAAATTGACCG 894  
868 GCGTGAAGTTGACCGGCTGCGGAGTACCGGTGCGAAGATTAAACTCAAGAAATTGACCG 927  
895 GGAACCCGCAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAGAACTTAACT 954  
928 GGGCCCGCAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAGAACTTAACT 987  
955 ACCCTTGAATGATGATTAATTCGATGATTAATTCGATGCAACGCGAAGAACTTAACT 1013  
988 GGTCTTGAATGATGATTAATTCGATGATTAATTCGATGCAACGCGAAGAACTTAACT 1047  
1014 GGTCTTGAATGATGATTAATTCGATGATTAATTCGATGCAACGCGAAGAACTTAACT 1073  
1048 GGTCTTGAATGATGATTAATTCGATGATTAATTCGATGCAACGCGAAGAACTTAACT 1107  
1074 GGTCTTGAATGATGATTAATTCGATGATTAATTCGATGCAACGCGAAGAACTTAACT 1132  
1108 GGTCTTGAATGATGATTAATTCGATGATTAATTCGATGCAACGCGAAGAACTTAACT 1167  
1133 ACMAACCGGAGGAGGATGAGTCAAGTCTCATGATGATGATGATGATGATGATGATGAT 1192  
1168 ATMAATCGAGGAGGATGAGTCAAGTCTCATGATGATGATGATGATGATGATGATGAT 1227  
1193 ACACGTAATCAATGAGCGGTGACAGAGGTTGCAACCGCGAGGAGGAGGATTAATCTCAG 1252  
1228 ACACGTAATCAATGAGCGGTGACAGAGGTTGCAACCGCGAGGAGGAGGATTAATCTCAG 1287  
1253 AAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312  
1288 AAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347  
1313 GTAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1372  
1348 GTAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1407  
1373 CACACCATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1432  
1408 CACACCATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1467  
1433 CACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1492  
1468 CACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1527

RESULT 46  
US-09-492-709A-242  
Sequence 242, Application US/09492709A  
GENERAL INFORMATION:  
APPLICANT: Zykind, Judith  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Trawick, John  
APPLICANT: Forsyth, R. Allyn  
APPLICANT: Freilich, Jamie M.  
APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN  
TITLE OF INVENTION: ESCHERICHIA COLI  
FILE REFERENCE: E1177A.001A  
CURRENT APPLICATION NUMBER: US/09/492,709A  
CURRENT FILING DATE: 2000-01-27  
NUMBER OF SEQ ID NOS: 485  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 242  
TYPE: DNA  
ORGANISM: E. Coli  
US-09-492-709A-242  
Query Match 67.5%; Score 984.2; DB 3; Length 1549;  
Best Local Similarity 82.2%; Pred. No. 0;  
Matches 1204; Conservative 0; Mismatches 253; Indels 8; Gaps 6;  
1 ATTGAACGCTGGCGGATGCTTTTACATGCAATGCAAGTCAAGGAGGAGGAGGAT--GCTTGC 58  
28 ATTGAACGCTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 87  
59 ATCT--GGTGGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 116  
88 TCTTCTGCTGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 147  
117 GGTAAACGATCAAGAAAGATGCTTAATACGATATATCTTAAGAGGAGGAGGAGGAGGAGGAGG 176  
148 GATTAACGATCAAGAAAGATGCTTAATACGATATATCTTAAGAGGAGGAGGAGGAGGAGGAGG 207  
177 GAAAGACCTTGGCGCTTTTGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 236  
208 TTTGGGCGCTTCTTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 267  
237 CTAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 296  
268 TCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 327  
297 CACGCGCCGAGCTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 356  
328 CACGCGCCGAGCTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 387  
357 GATTCAGCAATGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 416  
388 GATTCAGCAATGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 447  
417 AAGAAAGGTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 476  
448 AAGAAAGGTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 507  
477 TAACTACGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 536  
508 TAACTACGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 567  
537 GCGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 595  
568 GCGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 627  
596 GGAATTCGCTTTGAAACTCAAGGCTAGAGTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 655  
628 GGAATTCGCTTTGAAACTCAAGGCTAGAGTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 687  
656 AGCACTGAAATGCGTGAAGATTTGGAAGAACTGATGGCGAAGGAGGAGGAGGAGGAGGAGGAGG 715  
688 AGCGGTGAATGCGTGAAGATTTGGAAGAACTGATGGCGAAGGAGGAGGAGGAGGAGGAGGAGG 747  
716 ACATGAGCGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTAGTCC 775  
748 AGACTGAGCGCTCATGCAAGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTAGTCC 807  
776 AGCGCTTAAACGATGCACTGATTTGTTGGGCTTATTAGGCTTGG-TAACGAAGCTAAC 834

```
Db      808  ACGCCGTAACGATGTCGACTTGGAGTGTGTCCTTGAGCGGTGCTTCGGAGCTAAC  867
Qy      835  GCGTGAAGTTGACCGCCCTGGGAGTACGCTCCGAAGATTAAACCTCAAGAAATTGACGG  894
Db      868  GCGTTAAGTCGACCGCTGGGAGTACGCTCCGAAGTTAAACCTCAAGTAATGACGG  927
Qy      895  GGAACCCGCAAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTA  954
Db      928  GGGCCCGCAAGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTA  987
Qy      955  ACCCTTGACATGATGACGAATTTTCTAGAGAT-AGATTAGTCTTGGGAGCGCTAAC  1013
Db      988  GGTCTTGACATGACGAAGATTCTTCAAGATGAGATGCTTGGGAGCGCTAGACA  1047
Qy      1014  GGTCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  1073
Db      1048  GGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  1107
Qy      1074  CGCAACCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  1132
Db      1108  CGCAACCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  1167
Qy      1133  ACAAACCGGAGAGGTGGGAGTGAAGTCAAGTCTCATGCTGCTGCTGCTGCTGCT  1192
Db      1168  ATAACTGAGAGAGGTGGGAGTGAAGTCAAGTCTCATGCTGCTGCTGCTGCTGCT  1227
Qy      1193  ACACTTAATACAAATGGCGCTACAGAGGTTGSCAAACCGGAGGAGGAGCTAATCT  1252
Db      1228  ACACGTGCTACAAATGGCGCTACAGAGGTTGSCAAACCGGAGGAGGAGCTAATCT  1287
Qy      1253  AAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  1312
Db      1288  AAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  1347
Qy      1313  GTAATCGCGGATCAGCATGTGCGGTGAATACGTTCCGGGCTTGTACACACCGCG  1372
Db      1348  GTAATCGTGAATCAAGATGCGAGTGAATACGTTCCGGGCTTGTACACACCGCG  1407
Qy      1373  CACACCAATGGAGTGGTTCACAGAAAGAGTGTCTAACCGTAAGAGGCGCTTGC  1432
Db      1408  CACACCAATGGAGTGGTTCACAGAAAGAGTGTCTAACCGTAAGAGGCGCTTAC  1467
Qy      1433  CACGCTGATTCATGACTGGGTG  1457
Db      1468  CACTTGTGATTCATGACTGGGTG  1492

RESULT 47
US-09-492-709A-402
; Sequence 402, Application US/09492709A
; Patent No. 6720139
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; FILE REFERENCE: ESCHERICHIA COLI
; CURRENT APPLICATION NUMBER: US/09/492,709A
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 402
; LENGTH: 1549
; TYPE: RNA
; ORGANISM: E. Coli
US-09-492-709A-402
```

```
Query Match      67.5%; Score 984.2; DB 3; Length 1549;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 969; Conservative 235; Mismatches 253; Indels 8; Gaps 6;

Qy      1  ATTGAAGCTGGCGGATGCTTAAACATGCAAGTCCAGCGACGAGT--GCTTGC  58
Db      28  AATGAAGCTGGCGGATGCTTAAACATGCAAGTCCAGCGACGAGT--GCTTGC  87
Qy      59  ATCT--GGTGGCGATGGCGGATGCTTAAACATGCAAGTCCAGCGACGAGT  116
Db      88  UGCUUCGUAAGAGUGGCGGATGCTTAAACATGCAAGTCCAGCGACGAGT  147
Qy      117  GGTAAACGATGCAAGATGCTTAAACATGCAAGTCCAGCGACGAGTCCAG  176
Db      148  GAUAACUAACGUAACGUAACGUAACGUAACGUAACGUAACGUAACGUAAC  207
Qy      177  GAAAGACCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  236
Db      208  UUCGGGCGUUCGUAACGUAACGUAACGUAACGUAACGUAACGUAACGUA  267
Qy      237  CTACCAAGGCGAGATGCTTAAACATGCAAGTCCAGCGACGAGTCCAG  296
Db      268  UACCAAGGCGAGATGCTTAAACATGCAAGTCCAGCGACGAGTCCAG  327
Qy      297  CACGCGCGAGATGCTTAAACATGCAAGTCCAGCGACGAGTCCAG  356
Db      328  CACGCGCGAGATGCTTAAACATGCAAGTCCAGCGACGAGTCCAG  387
Qy      357  GATCAGCAATGCGCGGATGCTTAAACATGCAAGTCCAGCGACGAGT  416
Db      388  GAUCAGCAATGCGCGGATGCTTAAACATGCAAGTCCAGCGACGAGT  447
Qy      417  AAGAAAGCTTAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  476
Db      448  AAGAAAGCTTAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  507
Qy      477  TAACTAGTGCAGACGCGGATGCTTAAACATGCTGCTGCTGCTGCTGCT  536
Db      508  UAAUCUGGUGCAACGCGGATGCTTAAACATGCTGCTGCTGCTGCTGCT  567
Qy      537  GCGTAAAGGCTGCGGATGCTTAAACATGCTGCTGCTGCTGCTGCTGCT  595
Db      568  GCGTAAAGGCTGCGGATGCTTAAACATGCTGCTGCTGCTGCTGCTGCT  627
Qy      596  GGAATTCGTTGAACTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG  655
Db      628  GGAACUGAUCUGAUCUGAUCUGAUCUGAUCUGAUCUGAUCUGAUCUG  687
Qy      656  AGCAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  715
Db      688  AGCGGUAAGUGGUAAGUGGUAAGUGGUAAGUGGUAAGUGGUAAGUGG  747
Qy      716  AACTGACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTAATCTGCT  775
Db      748  AGACUGACGCTCATGCAAGAAAGCGTGGGAGCAACAGATTAATCTGCT  807
Qy      776  ACGCCTTAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  834
Db      808  ACGCCTTAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  867
Qy      835  GCGTGAAGTTGACCGCCCTGGGAGTACGCTCCGAAGATTAAACCTCA  894
Db      868  GCGTGAAGTTGACCGCCCTGGGAGTACGCTCCGAAGATTAAACCTCA  927
Qy      895  GGAACCCGCAAGCGGTGATTAATGATTAATGATTAATGATTAATGAT  954
Db      928  GGGCCCGCAAGCGGTGATTAATGATTAATGATTAATGATTAATGAT  987
Qy      955  ACCCTTGACATGATGACGAATTTTCTAGAGAT-AGATTAGTCTTGG  1013
Db      988  GGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  1047
Qy      1014  GGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  1073
```



QY 1328 CATGTCGGGTAAATACGTTCCCGGGTCTTGTACACACCGCCGTACACATCGGAGTG 1387  
DB 1316 AATGCCGCGGTAAATACGTTCCCGGGCCTTGTACACACCGCCGTACACATCGGAGTG 1375  
QY 1388 GGTTCACCAAGAGAGAGTGTCTTAACCGTAAAGAGGCGCTTGGCCACGGTGAG 1441  
DB 1376 GGTTCACCAAGAGAGTGTGTAAACCTTCGGAGGCGCTTACCATCTTG 1429

RESULT 49  
US-10-701-200-81  
; Sequence 81, Application US/10701200  
; Patent No. 6958222  
; GENERAL INFORMATION:  
; APPLICANT: Koffas, Matheos  
; APPLICANT: Odum, James M  
; APPLICANT: Schenize, Andreas J  
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN  
; FILE REFERENCE: CL1596 US NA  
; CURRENT APPLICATION NUMBER: US/10/701,200  
; CURRENT FILING DATE: 2003-11-04  
; PRIOR APPLICATION NUMBER: US/09/934,868  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/229,858  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 81  
; LENGTH: 1429  
; TYPE: DNA  
; ORGANISM: METHYLOMONAS SP.  
US-10-701-200-81

Query Match 67.4%; Score 982.4; DB 3; Length 1429;  
Best Local Similarity 83.5%; Pred. No. 0;  
Matches 1198; Conservative 0; Mismatches 226; Indels 10; Gaps 7;

QY 13 CGGCATGCTTTACATGCAATCGAAGCGAGCAGCGATGCTGATCTGGTGGAGT 72  
DB 1 CGGTATGCTTTACATGCAATCGAAGCGAGCAGCGATGCTGATCTGGTGGAGT 56  
QY 73 GCGCAGCGGAGTATGATCGAAGCGATCTGCAAGAGGAGGAGTAAACGATCGAAAG 132  
DB 57 GCGCAGCGGAGTATGATCGAAGCGATCTGCAAGAGGAGGAGTAAACGATCGAAAG 116  
QY 133 ATGCTCTAATACCGCATATCTCTAAGAGAGAAAGCAGGAGATCGAAGAGCTTGCGCTT 192  
DB 117 TCACGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGAGATCGAAGAGCTTGCGCTT 176  
QY 193 TTGGAGCGGCGATGCTGATTAAGTGTGGGTTAAAGCGCTACCAAGCGAGCAT 252  
DB 177 ATAGTATGAGCTTATGCTGATTAAGTGTGGGTTAAAGCGCTACCAAGCGAGCAT 236  
QY 253 CAGTATGCTGCTGAGAGAGCAGCAGCAGCTGAGACTGAGACAGCGCCAGACTCT 312  
DB 227 CCGTATGCTGCTGAGAGAGTATGAGCAGCAGCTGAGACTGAGACAGCGCCAGACTCT 296  
QY 313 ACGGAGGAGCAGTGGGGAATTTTGGACATGGGCGCAAGCTGATCCAGCAATGCCGC 372  
DB 297 ACGGAGGAGCAGTGGGGAATTTTGGACATGGGCGCAAGCTGATCCAGCAATGCCGC 356  
QY 373 GTGAGTGAAGAGCGCTTCGGGTTGTAAGCTCTTTCAGTCCAGGAAGAAAGTTACGGT 432  
DB 357 GTGAGTGAAGAGCGCTTCGGGTTGTAAGCTCTTTCAGTCCAGGAAGAAAGTTACGGT 416  
QY 433 AATATATGCTGCTATGATGAGCGGTATCGACAGAAAGCAGCGCTTATCATCTGCGACAG 492  
DB 417 TATATCCCGGTATGAC-TGACATTAACCATCAAGAGAGCAGCGCTTATCATCTGCGACAG 475  
QY 493 GCGCGGTTATATCGTAGGAGTGCAGCGTTATCGAATTTACTGGCGTAAAGGTTGCCA 552  
DB 476 GCGCGGTTATATCGTAGGAGTGCAGCGTTATCGAATTTACTGGCGTAAAGGTTGCCA 535

QY 553 GCGCGCTTTGTATAGTCAGATGTAATCCCGGGCTTAACCTGGGAATTCGCTTTGAAC 612  
DB 536 GCGCGCTTTTATATGATGATGTAAGCGCTTGGGCTTAACCTGGGAATTCGATTTGATAC 595  
QY 613 TACAAAGCTTATAGTGTGCAAGAGAGTGAATTTCCATGTGTAGCAATGAATTCGCTAG 672  
DB 596 TGGGAACTTATAGTGTGCAAGAGAGTGAATTTCCATGTGTAGCAATGAATTCGCTAG 655  
QY 673 AGATATGGAAGAACTGATGCGCAAGGAGCAGCTCTTGGGTTAACTGACGCTCATGCA 732  
DB 656 AGATCTGAAGAACTGATGCGCAAGGAGCAGCTCTTGGGTTAACTGACGCTCATGCA 715  
QY 733 CGAAAGCGTGGGAGCAAAACAGATTAGTACCTTGTAGTCCAGCGCTTAAAGATGTC 792  
DB 716 CGAAAGCGTGGTGTGCAAAACAGATTAGTACCTTGTAGTCCAGCGCTTAAAGATGTC 775  
QY 793 AACTATGTTGGG-CTTATATAGCTTGTGTAAGAGAGTGAACCGGTGAAGTTGACCGCC 851  
DB 776 AACTAAACGTTGGGTTCTTAAAGAACTTATAGTGTGAGCTTAAAGTGAACCGCC 835  
QY 852 TGGGAGTATGAGTGGCAAGATTAAACTCAAGGAATTGACGGGAGCCCGCAAGCGGT 911  
DB 836 TGGGAGTATGAGTGGCAAGATTAAACTCAAGGAATTGACGGGAGCCCGCAAGCGGT 895  
QY 912 GAATTATGATGATTATTCATGATGCAAGCGCAAAACCTTACCTTGAATGATGCG 971  
DB 896 GAAGATATGATGATTATTCATGATGCAAGCGCAAAACCTTACCTTGAATGATGCG 955  
QY 972 AATTTTCTAGATATGATTAGT-CTTGGGAAACGCTTAAACAGAGTGTGATGAGTGC 1030  
DB 956 AACTGTGACAGATGACTGTGCTTGGGAAACGAGAGCAGAGTGTGATGAGTGC 1015  
QY 1031 GTCAAGCTGCTGCTGAGATGTTGGGTTAACTCCCGCAAGAGCGCAACCTTGCATTA 1090  
DB 1016 GTCAAGCTGCTGCTGAGATGTTGGGTTAACTCCCGCAAGAGCGCAACCTTGCATTA 1075  
QY 1091 ATTGCA--TCATTTGTTGGGCACTTATATGAGACTCCCGTGAACAAACCGAGGAAG 1148  
DB 1076 GTTGCAGCGCGCTATGAGCGGGAACCTTATGAGAGACTCCCGTATTAACCGAGGAAG 1135  
QY 1149 TGGGATGAGCTCAAGTCTTATGAGCGCTTATGAGAGCGCTTATGAGAGCTTATGAG 1208  
DB 1136 TGGGATGAGCTCAAGTCTTATGAGCGCGCTTATGAGAGCGCTTATGAGAGCTTATGAG 1195  
QY 1209 GCGGTAAGAGGTTGGCAACCGCGAGGAGGAGTATCTCAAGAAAGCGCGTGTAGT 1268  
DB 1196 TCGGTAAGAGGTTGGCAACCGCGAGGAGGAGTATCTCAAGAAAGCGCGTGTAGT 1255  
QY 1269 CGATTCGAGTCTGCAACTGAC-TCCGTGAAGTGGAAATGCTATGATTCGCGATCAG 1327  
DB 1256 CGATTCGAGTCTGCAACTGAC-TCCGTGAAGTGGAAATGCTATGATTCGCGATCAG 1315  
QY 1328 CATGTCGGGTAAATACGTTCCCGGGTCTTGTACACACCGCCGTACACATCGGAGTG 1387  
DB 1316 AATGCCGCGGTAAATACGTTCCCGGGTCTTGTACACACCGCCGTACACATCGGAGTG 1375  
QY 1388 GGTTCACCAAGAGAGTGTCTTAACCGTAAAGAGGCGCTTGGCCACGGTGAG 1441  
DB 1376 GGTTCACCAAGAGAGTGTGTAAACCTTCGGAGGCGCTTACCATCTTG 1429

RESULT 50  
US-09-596-002-41/C  
; Sequence 41, Application US/09596002  
; Patent No. 6632636  
; GENERAL INFORMATION:  
; APPLICANT: Lagace, Robert, E.  
; APPLICANT: Paterson, Chandra  
; APPLICANT: Berg, Kim, L.  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME  
; FILE REFERENCE: PM-0008-4 US  
; CURRENT APPLICATION NUMBER: US/09/596,002  
; CURRENT FILING DATE: 2000-06-16



```

; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 41
; LENGTH: 269223
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incycle template ID No. 6632636 41
; PUBLICATION INFORMATION:
US-09-596-002-41

Query Match      67.1%; Score 977.4; DB 3; Length 269223;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 1208; Conservative 0; Mismatches 241; Indels 16; Gaps 7;

QY      1  ATTGAACGCTGCGCGCATGCTTTACACATGCAAGTCGAACGCGCAGCAGCGATGCTTGCAAT 60
Db      92966  ATTGAACGCTGCGCGCATGCTTTACACATGCAAGTCGAACGA-AGTTAGGAAGCTTGCTT 92908

QY      61  CTGGGCGAGTGGGCGACGGGTGAGTAATGATGGAACGTAATCCAGAGAGGGGGGTA 120
Db      92907  CTGATACCTTAGTGGCGACGGGTGAGTAATGCTTAGGAATCTGCTTAGTGGGGGTA 92848

QY      121  ACGCATCGAAAGATGTGCTAATACCGCATATCTCTAAGAGAAAGCAGGGGATCGAA 180
Db      92847  ACTTGGGAAACCCAGCTAATACCGCATACGACTACGGGTGAAGGGGGCTTT----- 92793

QY      181  GACCTTGGCGCTTTTGAAGCGCGCATGCTGTAGTAAGTGAAGTGGGTGAAGGCTTAC 240
Db      92792  TAGCTCTGCTATTTAGATGAGCTTAAGTCGATTAAGTGTGGTGAAGGCTTAC 92733

QY      241  CAAGCGCATGATCAATGTTGCTGAGAGCAGACAGCCACACTGGGACTGAGACAGC 300
Db      92732  CAAGCGCATGATCTTGAAGTGTGAGAGATGATCAGCCACACTGGGACTGAGACAGC 92673

QY      301  GCCCAGACTCTTACGCGAGGAGCAGAGTGGGGGAAT-----TTTGAACAATGGCGCAAGCC 355
Db      92672  GCCCAGACTCTTACGCGAGGAGCAGAGTGGGGGAATTTGGAATTGGAACAATGGCGCAAGCC 92613

QY      356  TGATCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGTTGAAGCTCTTTCAGTCA 415
Db      92612  TGATCAGCAATGCGCGGTGAGTGAAGAGGCTTGGTGAAGCACTTTAAGTGGG 92553

QY      416  GAAAGAAAGTTACGTTAATATGCTGACTCATGACGCTATCGACAGAGAAAGCACCGG 475
Db      92552  GAGGAAAGCTTATGTTAATACCCATTAAGCCCTGACCTTACCAAGATTAAGCACCGG 92493

QY      476  CTAACCTAGCTGCAGACGCGCGGTATACGTAAGGTGCAAGGCTTAATCGGAATTACTG 535
Db      92492  CTAACCTAGCTGCAGACGCGCGGTATACGTAAGGTGCAAGGCTTAATCGGAATTACTG 92433

QY      536  GCGCTAAGAGGTGCGCAGCGCGCTTGTGAAGTCAGATGTGAATCCCGCGGCTTAACCTG 595
Db      92432  GCGCTAAGAGGTGCGCAGCGCGCTTGTGAAGTCAGATGTGAATCCCGCGGCTTAACCTG 92373

QY      596  GGAATTGCGTTGAACCTACAGAGCTTAGAGTGTGCGCAGAGGAGGTGGAATTCATGTGT 655
Db      92372  GGAATGCTGATCTGATCTGATTAATCTAGAGTGAAGTGAAGGAGGTGAATTCAGAGTGT 92313

QY      656  ACGAGTGAAGTGTGAGATATGAAGAACATGATGGCGAAGGAGCGCTCTGGGTTA 715
Db      92312  ACGAGTGAAGTGTGAGATATCTGAGAGTAATCCGATGGCGAAGGAGCGCTCTGGCATT 92253

QY      716  ACACTGACGCTCATGACGAAACGCTGGGAGGAAACAGGATTAGATACCTTGTAATCC 775
Db      92252  ATACTGACACTGAGGTGCGAAACGCTGGGTAAGCAAGATTAAGTACCTTGTAATCC 92193

QY      776  AGCGCTTAACGATGTCACTAATGTTGTGGGCTTAT-TAGGCTTGGTAAAGCAAGCTAAC 834
Db      92192  AGCGCTTAACGATGTCTACCAATCGTTGGGCTTTTAAAGACTTAGTGAAGCAAGTTAAC 92133
```

```

QY      835  GCGTGAAGTTGACCGGCTGGGGAGTAGAGGTGCAAGATTAAACTCAAAGGAATTGACGG 894
Db      92132  GCAATTAAGTAAACCGGCTGGGGAGTAGAGGTGCAAGATTAAACTCAAATGAATTGACGG 92073

QY      895  GGACCCGCAAGAGGGGTGATATGATGATTAATTCGATGCAAGCGGAAACCTTACTT 954
Db      92072  GGGCCCGCAAGAGGGGTGATGATGATTAATTCGATGCAAGCGGAAACCTTACTT 92013

QY      955  ACCCTGACATGAGCGAATTTCTAGATTAATTAATG-CTTCCGGGAAGCTTAACCA 1013
Db      92012  GATCTTGAATAGAGAAATCTTCAAGATGCAAGAGTGCCTTCGGGAATTCATATCA 91953

QY      1014  GGTGTGATGAGCTGTCTGACGCTCGTGTGCGTGAAGTGTGGGTTAAGTCCGCAACGAG 1073
Db      91952  GGTGTGATGAGCTGTCTGACGCTCGTGTGCGTGAAGTGTGGGTTAAGTCCGCAACGAG 91893

QY      1074  CGCAACCTTGTCAATTAATGCAATC-ATTGGTTGGGCACTTTAATGAGACTGCGGTTG 1132
Db      91892  CGCAACCTTGTCTTAATGCAATCAGCAGCTCGTGGGAACTTAAAGATTAATGCAAGT 91833

QY      1133  ACAAACCGAGAGAGTGGGATGACGTAAGTCTTCAATGACCTTATGGGTAGGGCTTC 1192
Db      91832  ACAAACCGAGAGAGTGGGATGACGTAAGTCTTCAATGACCTTATGACCGAGGCTTAC 91773

QY      1193  ACAGTATTAATGAGCGCGCTACAGAGGTTGCCAACCGGAGGGGAGCTAATCTGAG 1252
Db      91772  ACAGTGTCTACAAATGTTGTGTAACAAAGGTTGCTACAGAGATGATGCTTAATCTCA 91713

QY      1253  AAGCGCGCTGATGCGGATCGAGTCTGCAACTGCACTCCGTGAAGTGGGAATCGCTA 1312
Db      91712  AAGCGCAATGTTAGTCCGATTTGAGTCTGCAACTGCACTCCATGAAGTGGGAATCGCTA 91653

QY      1313  GTAATCGGATACAGCATGTGCGGTGAATAGTTCCCGGCTTGTACACACCGCGCT 1372
Db      91652  GTAATCGGATACAGCATGTGCGGTGAATAGTTCCCGGCTTGTACACACCGCGCT 91593

QY      1373  CACACATGAGAGTGGGTTTCAACCAAGCAGTGTGTAACCGTAAGGAGGGCGCTTGC 1432
Db      91592  CACACATGAGAGTGTGATCTCAACCAAGTGTGTAACCGTAAGGAGGGCGATCAC 91535

QY      1433  CACGCTGAGATTCAATGACTGGGTTG 1457
Db      91534  CACGCTGAGATTCAATGACTGGGTTG 91510

RESULT 51
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: 08/476,102  
FILING DATE: JUN-5-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41, 971  
REFERENCE/DOCKET NUMBER: PB186P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-557-884-1

Query Match 66.7%; Score 971.8; DB 3; Length 1830121;  
Best Local Similarity 81.3%; Pred. No. 0; Mismatches 267; Indels 7; Gaps 5;  
Matches 1189; Conservative 0;

1 ATTGAACGCTGCGGCATGCTTTTACATGCAAGTCGAACGCGCAGCAGCG--ATGCTTGC 58  
770639 ATTGAACGCTGCGGCATGCTTTTACATGCAAGTCGAACGCGTACGAGAAAGCTTGC 770698  
59 ATC--TGTGCGAGTGGCGGACGCGGTGATGATGCAATCGGAACCTATCCAGAGAGGG 116  
770699 TTTCTGTGACGAGTGGCGGACGCGGTGATGATGCTTGGGAATCGCTTATGAGAGGG 770758  
117 GGTAAACGATGGAAGATGTGTATATCCGATATACCTCTAAGGAGGAAAGCGGAGATC 176  
770759 GATTAACGAGGAAACCTGTCTGTATATACCGCTATATCGGAAGTGAAGTGCAGGACT 770818  
177 GAAAGACCTTGGCTTTTGAAGCGGCGCATGCTGATTAGCTAGTGGTGGGTAAAGGC 236  
770819 GAGAGCGCGCATGCGCATAGATGAGCCAGTGGGATTAAGTATGGTGGGTAAAGGC 770878  
237 CTACCAAGCGACGATCGATGATGTTGTTGAGAGGACGACCGACCACTGSGAATTGAGA 296  
770879 CTACCAAGCGTGCATCTCTAGCTGATGAGAGGATGACCGACCACTGGAATCGAGA 770938  
297 CACGCGCCAGACTCTTAAGGAGGAGGACGAGTGGGGAATTTTGAAGTGGGCGGAGGCT 356  
770939 CACGCTCCAGACTCTTAAGGAGGAGGACGAGTGGGGAATTTGCGCAATGGGCGGAGCCT 770998  
357 GATCCAGCAATGCGCGTGAAGTGAAGAGGCGCTTGGGTTGTAAGCTCTTTCAGTCCAG 416  
770999 GACGACGCAATGCGCGTGAAGTGAAGAGGCGCTTGGGTTGTAAGCTCTTTCAGTATGG 771058  
417 AAGAAAGGTTACGCTAATATCTGACTCATGACGCTATCGACAGAAAGAACCGGC 476  
771059 AAGAAAGGTTGATGTGTATATAGTATCATCAATTAATGACGTTAATACAGAAAGCACCGGC 771118  
477 TAACTACGTGCCAGAGCGCGGTAAATCGTAGGGTGAAGGCGTAAATCGGAATTAACGG 536  
771119 TAACTACGTGCCAGAGCGCGGTAAATCGTAGGGTGAAGGCGTAAATCGGAATTAACGG 771178  
537 GCGTAAAGGTCGCGAGCGCGCTTGTAAAGTCAATGTGAATATCCCGGCGCTTAACCTGG 596  
771179 GCGTAAAGGTCGCGAGCGCGCTTGTAAAGTCAATGTGAATATCCCGGCGCTTAACCTGG 771238  
597 GAATTCGCTTGAATCTAACAGGCTAAGTGTGCGAGAGGAGGTGAATTCATGTGTGA 656  
771239 GAATTCGCTTGAATCTAACAGGCTAAGTGTGCGAGAGGAGGTGAATTCATGTGTGA 771298  
657 GCAAGTGAATGCTGATGATATGAGAAACATCGATGGGAGGAGGACCTCTCGGTTAA 716  
771299 GCGGTGAATGCTGATGATATGAGAAACATCGATGGGAGGAGGACCTCTCGGTTAA 771358  
717 CACTGACGCTGATGACGAAAGCGTGGGAGCAACAGGATTAAGTAACTCTGATGTCGA 776  
771359 TACTGACGCTGATGACGAAAGCGTGGGAGCAACAGGATTAAGTAACTCTGATGTCGA 771418

777 CGCCCTAAAGATGTCACCTAGTGTGGGCTTATTAGGCTTGTAACGAGCTAACGC 836  
771419 CGCTGTAAACGCTTTCGATTTTGGGATTTGGGCTTA--GAGCTTGGTGGCCGTAACGTA 771477  
837 GTGAAGTGAACCGCTGGGAGTACCGTTCGCAAGATTAACTCAAGAAATTAACGGGG 896  
771478 GATTAATGACACCGCTGGGAGTACCGGCGGCAAGTTAAACTCAATGAATTAACGGGG 771537  
897 ACCCGCACAGCGGTGATTTATGAGATTAATTTGATGCAACGGGAAACCTTAACCTAC 956  
771538 GCCCGCACAGCGGTGAGCATGTGTTAATTTGATGCAACGGGAAACCTTAACCTAC 771597  
957 CTTGACATGTAAGCAATTTTCTAGAGAT--AGATTAGTCTTGGGAAACGCTTAACACAG 1015  
771598 TCTTGACATCTTAAGAAAGCTCAGAGATGCTTGGCTTCCGGAACTTAAGACAGG 771657  
1016 TGCTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075  
771658 TGCTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771717  
1076 CAACCTTGTATTATTTGCTATC--ATTGGTTGGGACCTTAATGAGACTGCGGTAC 1134  
771718 CAACCTTATCTTGTGTTGCTGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771777  
1135 AAACCGGAGAAAGGAGGAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1194  
771778 AAACCGGAGAAAGGAGGAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 771837  
1195 ACGTATATCAATGAGCGCTGACAGAGGTTGCAACCGCGAGGAGGAGTCAATTCAGAA 1254  
771838 ACGTATATCAATGAGCGCTGACAGAGGTTGCAACCGCGAGGAGGAGTCAATTCAGAA 771897  
1255 AGCGGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1314  
771898 AGTACGTCTAAGTCCGATGAGGAGTCTGCAACTGACCTCAATGAAGTCAAGTCAAGT 771957  
1315 AATGCGGATGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1374  
771958 AATGCGGATGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 772017  
1375 CACCATGAGAGTGGGTTTCAACAGAGGAGTGTACTTAACCGTAAGAGGCGCTTGCCA 1434  
772018 CACCATGAGAGTGGGTTTCAACAGAGGAGTGTACTTAACCGTAAGAGGCGCTTGCCA 772077  
1435 CGGTGAGTTCAATGACTGGGGTG 1457  
772078 CGGTGAGTTCAATGACTGGGGTG 772100

RESULT 52  
US-09-557-884-1/c  
Sequence 1, Application US/09557884  
Patent No. 6506581  
GENERAL INFORMATION:  
Applicant: Fleischmann et al.  
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses thereof  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884  
 FILING DATE: 25-Apr-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/476,102  
 FILING DATE: JUN-5-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Michelle S. Marks  
 REGISTRATION NUMBER: 41,971  
 REFERENCE/DOCKET NUMBER: PB186P3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301-309-8504  
 TELEFAX: 301-309-8439  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1830121 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-557-884-1  
 Query Match 66.7%; Score 971.8; DB 3; Length 1830121;  
 Best Local Similarity 81.3%; Pred. No. 0;  
 Matches 1189; Conservative 0; Mismatches 267; Indels 7; Gaps 5;  
 QY 1 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAACGCGACGCG--ATGCTTCC 58  
 DB 128095 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAACGCGACGCG--ATGCTTCC 128036  
 QY 59 ATC--TGATGCGAGTGGCGGAGCGGAGTGAATGATCGGAACGTATCCAGAGAGGGG 116  
 DB 128035 TTTCTTGCTGAGAGAGTGGCGGAGCGGAGTGAATGCTTGGGAATCGCTTAATGAGGGG 127976  
 QY 117 GGTAAACGATCGAAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATC 176  
 DB 127975 GATTAACGAGGGGAAATGTCTGCTAATACCGCATATCTCTAAGAGAGAAAGTGGGAGCT 127916  
 QY 177 GAAAGACCTTGGCGCTTTTGGAGCGCGGATGCTGATGCTAGCTAGCTAGTGGGTTAAAGCC 236  
 DB 127915 GAGAGCGCCGATGCTAGTGAAGAGAGCCCAAGTGGGATGAGTGGTGGGTTAAATGCG 127856  
 QY 237 CTACCAAGCGAGCATCAGTATGTTGCTGAGAGAGAGCAGCAGCCACACTGGAGCTGAGA 296  
 DB 127855 CTACCAAGCGCTGAGATCTCTAGCTGGTCTGAGAGAGATGACCAAGCCACACTGGAACTGAGA 127796  
 QY 297 CACGCGCCAGATCTCTACGAGAGAGCGCAGTGGGGAAATTTTGAACAATGGGCCCAAGCCT 356  
 DB 127795 CACGCTCCAGACTCTTACGAGAGAGCGCAGTGGGGAAATTTTGGCAATGGGGGAAACCTT 127736  
 QY 357 GATCCAGCAATGCGCGGTGAGTGAAGAGGCTTGGGGTTTGAAGCTCTTTTCACTGAG 416  
 DB 127735 GACGCGAGCCATGCGCGGTGATGAAGAGGCTTGGGGTTTGAAGCTCTTTTGGGATTTG 127676  
 QY 417 AAGAAAGGTTACGTTAATATCTGACTCATGACCGTATCGACAGAGAGACACCGCG 476  
 DB 127675 AGAAGAGTTGATGTGTTAATAGTACATCAATTTGAGCTTAATACAGAGAGACACCGCG 127616  
 QY 477 TAACTACGTGCGAGCAGCGCGGTAAATAGTAGGTCAGAGCGTTAATCGAAATTACTGG 536  
 DB 127615 TAACTCCGTGCGAGCAGCGCGGTAAATAGTAGGTCAGAGCGTTAATCGAAATTACTGG 127556  
 QY 537 GCGTAAAGGTGGCGAGCGCGCTTTGTAAGTCAGATGTAATCCCGGGCTTAACCTGG 596  
 DB 127555 GCGTAAAGGTGGCGAGCGCGCTTTGTAAGTCAGATGTAATCCCGGGCTTAACCTGG 127496  
 QY 597 GAATGCGTTTGAACCTACAGGCTAGAGTGGCGAGAGGAGGAGTGAATTCATGTGTA 656  
 DB 127495 GAATGCGTTTGAACCTAGAGTGGCGAGAGGAGGAGTGAATTCATGTGTA 127436  
 QY 657 GCAAGTGAATGCGTGAAGATATGAGAGAACATGATGGCGAGAGCGACTTCCTGGGTTAA 716  
 DB 127435 GCGGTGAATGCGTGAAGATATGAGAGAACATGAGAGAGCGAGCGCTTCCTGGGATG 127376

QY 717 CACTGACGCTGATGACGAAAGCGTGGGAGCAAAACAGATTAAGTACCCTGTAAGTCCA 776  
 DB 127375 TACTGACGCTGATGACGAAAGCGTGGGAGCAAAACAGATTAAGTACCCTGTAAGTCCA 127316  
 QY 777 GCGCCTAAACAGATGCACTAGTGTGGGCTTATTAAGCTTGTGTAACGAAGCTTAACGC 836  
 DB 127315 CCGTGTAAACCGCTGTGCAATTTGGGGATTTGGGCTTA-GAGCTTGGTGGCCGTAGTAAAGT 127257  
 QY 837 GTGAAGTTGACCGCTGGGAGTACGATGTCGCAAGTTAAACTCAAAAGAAATTGACGGGG 896  
 DB 127256 GATAAATGACCGCTGGGAGTACGATGTCGCAAGTTAAACTCAAAAGAAATTGACGGGG 127197  
 QY 897 ACCCGCAACCGGCTGATTAATGAGTAAATTCGATGCAACGCAAAACCTTACCTAC 956  
 DB 127196 GCCCGCAACCGGCTGATTAATGAGTAAATTCGATGCAACGCAAAACCTTACCTAC 127137  
 QY 957 CCTTGACATGTAAGCAATTTTCTAGAGAT-AGATTAGTGTCTTGGGAAACGCTAACACAG 1015  
 DB 127136 TCTTGACATCTTAAGAAAGCTCAGAGATGAGCTTGTGCTTGGGAACTTAAGACACAG 127077  
 QY 1016 TCGTCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075  
 DB 127076 TCGTCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 127017  
 QY 1076 CAACCTTGTCTTAATTTGCCATC-ATTGCTTGGGCACTTTAATGAGACTGCGGCTGAC 1134  
 DB 127016 CAACCTTGTCTTAATTTGCCATC-ATTGCTTGGGCACTTTAATGAGACTGCGGCTGAC 126957  
 QY 1135 AAACGGAGGAAGTGGGAGTACGTCATGCTCCTATGAGGCTTAC 1194  
 DB 126956 AAACGGAGGAAGTGGGAGTACGTCATGCTCCTATGAGGCTTAC 126897  
 QY 1195 ACGTAAATCAATGCGCGCTACAGAGGTTGCCAACCCCGAGGGAGCTTAATCTGAA 1254  
 DB 126896 ACGTAAATCAATGCGCGCTACAGAGGTTGCCAACCCCGAGGGAGCTTAATCTGAA 126837  
 QY 1255 AGCGGCTGTAATCCGAGATCGAGATGCTCAATCCGAGTCCGTAAGTCCGTAGT 1314  
 DB 126836 AGTACGCTTAATCCGAGATGCTCAATCCGAGTCCGTAAGTCCGTAGT 126777  
 QY 1315 AATCGGATCAATGCTGCGGCTGATACGTTCCCGGCTTGTACACACCGCCGCTCA 1374  
 DB 126776 AATCGGATCAATGCTGCGGCTGATACGTTCCCGGCTTGTACACACCGCCGCTCA 126717  
 QY 1375 CACATGAGAGTGGGTTTACCAAGAGCAGTATGCTTAACGTTAAGAGAGGCGCTTGCA 1434  
 DB 126716 CACATGAGAGTGGGTTTACCAAGAGCAGTATGCTTAACGTTTGAAGGCGCTTACCA 126657  
 QY 1435 CCGTGAATTCATGACTGGGGTG 1457  
 DB 126656 CCGTGAATTCATGACTGGGGTG 126634  
 RESULT 53  
 US-09-643-990A-1  
 Sequence 1, Application US/09643990A  
 Patent No. 6528289  
 GENERAL INFORMATION:  
 APPLICANT: Robert D. Fleischmann  
 Mark D. Adams  
 Owen White  
 Hamilton O. Smith  
 J. Craig Venter  
 TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof  
 NUMBER OF SEQUENCES: 1  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville,  
 STATE: MD

```

COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-980A-1

Query Match      66.7%; Score 971.8; DB 3; Length 1830121;
Best Local Similarity 81.3%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 267; Indels 7; Gaps 5;

QY 1 ATGGAAGCTGGCGGCGATGCTTACACATGCAAGTCGAACGGCAGCAGCG--ATGCTTGC 58
DB 770639 ATGGAAGCTGGCGGCGGAGCTTACACATGCAAGTCGAACGGTACGAGAGAAAGCTTGC 770698
QY 59 ATC--TGCTGGCGAGTGGCGGAGCGGAGTGAATGATCATCGAACCCTATCCAGAAAGAGGG 116
DB 770699 TTTCTGTGAGAGATGGCGGAGCGGAGTGAATGCTTGGGAATCTGGCTTAATGAGAGGG 770758
QY 117 GATTAACGATGGAAGATGTGCTAATACCGCATTAATCTTAAGAGAAAGAGAGAGGGATC 176
DB 770759 GATTAACGAGGAGAACTGTCTGCTAATACCGCATTAATCGGAAGATGAAGTGCAGGACT 770818
QY 177 GAAGAAGCTTGGCGCTTTGGAGCGGCGATGCTGATTAGCTAATGTTGGGTGAAGGC 236
DB 770819 GAGAGGCGCATGCGATGAGATGAGAGCCCAAGTGGGATTAAGTGGTGGGTAAATGC 770878
QY 237 CTACCAAGGCGAGCATGATGATGTTGCTGAGAGAGCAGCAGCCACACTGGAGCTGAGA 296
DB 770879 CTACCAAGCTGCGATCTCTAGCTGTCTGAGAGAGTGAACAGCAGCACTGGAACTGAGA 770938
QY 297 CACGGCCCAAGACTCTTAAGGAGGAGCAGCAGATGEGGAAATTTTGAACAATGGGCGCAAGCT 356
DB 770939 CACGGTCCAGACTCTTAAGGAGGAGCAGCAGATGEGGAAATTTTGAACAATGGGCGGAGCT 770998
QY 357 GATCAGCAATGGCGCGAGTGAAGAGGCGCTTGCGGTTGTAAGCTCTTCACTGAG 416
DB 770999 GACGAGCAGTGGCGCGTGAATGAAGAGGCGCTTGCGGTTGTAAGCTCTTCACTGAG 771058
QY 417 AAGAAAGGTTAAGCTTAATATCGTGACTCATGACGGTATGAGAGAGAGAGAGCAGCCGC 476
DB 771059 AAGAAAGGTTAAGCTTAATATGATCATGAATTAAGCTTAATTAAGAGAGAGAGCAGCCGC 771118
QY 477 TAACTAGCTGCCAGAGCGCGGCTTAATACGTAAGGCTGCAAGGCTTAATTCGAAATTAAGT 536
DB 771119 TAACTAGCTGCCAGAGCGCGGCTTAATACGTAAGGCTGCAAGGCTTAATTCGAAATTAAGT 771178
QY 537 GCGTAAAGGAGTGGCGAGGCGGCTTTGTAAGTGAATGTGAATCCCGGCGCTTAACCTGG 596

```

```

DB 771179 GCGTAAAGGAGCAGCAGCGGCTTTAATTAAGTGAAGTGAAGAGCCCTGGGCTTAACCTAG 771238
QY 597 GAATTCGCTTTGAAACTTAACAAGCTTAAGATGTGGCAGAGGAGGTGAAATTCATGTGTA 656
DB 771239 GAATTCGATTTTCAAGCTGGTAACTAAGTACTTTAGGAGAGGGGTAAATTCACGTGTA 771298
QY 657 GCAGTGAATGCGTGAAGATTAATGAAGAACATCGATGCGGAGAGGAGGAGCTCTGGGTAA 716
DB 771299 GCGGTGAATGCGTGAAGATGTGAAGAAATCCAGAGCGGAGAGAGCCCTTGGGAATG 771358
QY 717 CACTGACGCTCATGACAGAAAGCGTGGGAGCAACAGATTAGTACCTGTAGTCCA 776
DB 771359 TACTGACGCTCATGAGGAAAGCGTGGGAGCAACAGATTAGTACCTGTAGTCCA 771418
QY 777 CGCCCTTAACGATGCTCAATGTTGTTGGCTTTATTAAGCTTGAACGAAGCTTAACGC 836
DB 771419 CGCTGAACGCTGCTCAATTTGGGAGATTGGGCTTA--GAGCTTGGTGCCTGAGCTTAACGT 771477
QY 837 GTGAAGTTGACCGGCTGGGAGTACGGTCCGAAGATTAACTCAAGAAATTGACGGGG 896
DB 771478 GATTAATTCAGACCGCTGGGAGTACGGCCGCAAGGTTAAACTCAATGAATTAACGGGG 771537
QY 897 ACCGCAAGCGGCTGATTTATGATTAATTCGATGCAACGCAAAACCTTAACCTAC 956
DB 771538 GCCGCAAGCGGCTGGAGCATGTGTTAATTCGATGCAACGCAAGAACTTAACCTAC 771597
QY 957 CTTGACATGTAAGCAATTTTCTAGAGT--AGATTAGCTCTTGGGAAACGCTTAACACAG 1015
DB 771598 TCTTGACATCTTAAGAAAGCTCAGAGTGAAGCTTGGCTTGGGAACTTAAGACAGG 771657
QY 1016 TGCTGACATGCTGACGTCAGCTGCTGTCGATGATGTTGGGTAAAGTCCGCAACGAGCG 1075
DB 771658 TGCTGACATGCTGTCAGCTGCTGTCGATGATGTTGGGTAAAGTCCGCAACGAGCG 771717
QY 1076 CAACCTTGTCAATTAATTTGCCATC--ATTGGTTGGGCACTTAATGAAGATCCGCTGAC 1134
DB 771718 CAACCTTATCTTTTGTGTCAGAGGACTTGGTCCGGAACCTCAAGAGAGATGCGCAGTAT 771777
QY 1135 AAACCGGAGAGGTTGGGATGACGTCATGCTCAATGAGCCCTTAATGGTATGGCTTAC 1194
DB 771778 AAACCTGAGAGAGGTTGGGATGACGTCATGAGCCCTTAATGAGTGGCTTAC 771837
QY 1195 ACGTATTAACAATGGCGGCTACAGAGGGTGGCAACCGCGAGGGAGCTAATCTCAGAA 1254
DB 771838 ACGTCTAACAATGGCTAATCAAGAGGAGAGGAGGAGCTGAGAGCTCAATCAATTA 771897
QY 1255 AGCGCTGTAGTCCGATCGGATCTGCAACTGCACTCGTGAAGTCGGAATGCTAGT 1314
DB 771898 AGTACGTCTAAGTCCGATTTGGAGTCTGCAACTGCACTCGATGAAGTCGGAATGCTAGT 771957
QY 1315 AATGCGGATGACATGTCGGGTGAATAGTTCCTCGGCTTTGTAACAACCGCCGCTCA 1374
DB 771958 AATGCGGATGACATGTCGGGTGAATAGTTCCTCGGCTTTGTAACAACCGCCGCTCA 772017
QY 1375 CACCATGGAGTGGGTTTCAACAGAGAGGTAGTCTTAACCGTAAAGAGGCGCTTGCA 1434
DB 772018 CACCATGGAGTGGGTTTCAACAGAGAGGTAGTCTTAACCTTTTGAAGGCGCTTAACA 772077
QY 1435 CGGTGAGATTGATGACTGGGCTG 1457
DB 772078 CGGTGATGATTGATGACTGGGCTG 772100

RESULT 54
US-09-643-990A-1/c
: Sequence 1, Application US/09643990A
: Patent No. 6528289
: GENERAL INFORMATION:
: APPLICANT: Robert D. Fleischmann
: Mark D. Adams
: Owen White
: Hamilton O. Smith
:

```

J. Craig Venter  
TITLE OF INVENTION: The Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville,  
STATE: MD  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/643,990A  
FILING DATE: 23-Aug-2000  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,429  
FILING DATE: 1995-06-07  
APPLICATION NUMBER: 08/426,787  
FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB186P1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-610-5790  
TELEFAX: 310-309-8439

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-643-990A-1

Query Match 66.7%; Score 971.8; DB 3; Length 1830121;  
Best Local Similarity 81.3%; Pred. No. 0;  
Matches 1189; Conservative 0; Mismatches 267; Indels 7; Gaps 5;

QY 1 ATTGAACGCTGCGGCGATCTTTACATGCAAGTGAACGCGACACG--ATGCTTGC 58  
DB 128095 ATTGAACGCTGCGGCGGCACTTAAACATGCAAGTGAACGCGACGGAAGGCTTGC 128036

QY 59 ATC--TGTGTGCGAAGTGGCGGCGGATGATGATGCGGAACGATCCAGAAAGGGG 116  
DB 128035 TTTCTTGTGACGAGTGGCGGAGGAGTGAATGCTTGGAAATCTGGCTTATGAGAGGG 127976

QY 117 GGTAAACGATCGAAGATGTGCTAATATACGATATATCTTAAGAGAGAAAGCAGGGATC 176  
DB 127975 GATTAACGAGGAAATGTGCTTAATATACGCGTATATATCGGAAGATGAAGATCGGGGATC 127916

QY 177 GAAAGACCTTGGCGCTTTTGAAGCGCGGATGCTGATTAAGTAACTGTTGGGTAAAGGC 236  
DB 127915 GAGAGCGCGCATGCTAATAGATGAGCCAAAGTGGGATTAAGTGTGGGTAAATATG 127856

QY 237 CTACCAAGCGACGATCAGATGTTGCTGAGAGAGAGCAGACCACTGGGAGCTGAG 296  
DB 127855 CTACCAAGCGCTGAGATCTTAGCTGTGCTGAGAGAGATGACAGCACACTGGAATGAGA 127796

QY 297 CACGCGCCAGACTCTTAACGGGAGCAGACAGTGGGGAATTTTGAACATGGGCGCAAGCCT 356  
DB 127795 CACGCTCCAGACTCTTAACGGGAGCAGACAGTGGGGAATTTTGGGCAATGGGGGAACCT 127736

QY 357 GATCCAGCAATGCGCGGTGAGTAAGAAAGCCTTGGGTTTGAAGCTCTTTCACTGAG 416  
DB 127735 GACGCAAGCCATGCGCGGTGAGTAAGAAAGCCTTGGGTTTGAAGTCTTTCTGATATG 127676

QY 417 AAGAAAAGTTACGCTAATTAATCGTGACTCATGACGGTATCGACAGAAAGACCGGC 476  
DB 127675 AAGAAAGTTGATGTGTTAATAGTACATCAATTAACGTTAATACAGAAAGACCGGC 127616

QY 477 TAACCTAGTGCACAGCGCGGTAAATACGTAAGGTGCGACGTTAATCGGAATTAATCTGG 536  
DB 127615 TAACCTCGTGCACAGCGCGGTAAATACGTAAGGTGCGACGTTAATCGGAATTAATCTGG 127556

QY 537 GCGTAAAGGTGCGACGCGGCTTTGATGTCAGATGTGAATCCCGGGCTTAACCTGG 596  
DB 127555 GCGTAAAGGACGACGCGGCTTTAATGAGAGGTGAAAGCCCTGGGCTTAACCTAG 127496

QY 597 GAATTCGCTTTGAACATCAAGCTAAGGTGCGACAGGAGGAGTGAATTCATGTGTA 656  
DB 127495 GAATTCGATTTCAAGCTGGGTAACTAGATGACTTTTGGAGAGGTAGTAATTCACGTGTA 127436

QY 657 GCACTGAATGCTGATAGATATGAGAAACATCGATGCGAAGGCGACGCTCTGGGTTAA 716  
DB 127435 GCGGTGAATGCTGATAGATATGAGAAATCCGAAAGCGAAGGCGACGCTCTGGGAATG 127376

QY 717 CACTGACGCTGATGACGAAAGCGTGGGAGCAACAGATTAATACCTGCTGTAGTCCA 776  
DB 127375 TACTGACGCTCATGTGCGAAAGCGTGGGAGCAACAGATTAATACCTGCTGTAGTCCA 127316

QY 777 CGCCTTAACGATGCTCACTAGTTGTTGGCCTTATTTAGGCTTGAAGAACTTAACGC 836  
DB 127315 CGCTGTAACGCTGCTGATTTGGGATTTGGGCTTA-GAGCTTGGTCCCTGATGATGCT 127257

QY 837 GTGAATGTAACGCGCTGGGAGTACGCTCGCAAGATTAACCAAGGAATTAACGCGG 896  
DB 127256 GATTAATGTAACGCTGGGAGTACGCGCGCAAGTAAATCAATTAATTAACGCGG 127197

QY 897 ACCCGCAAGCGGTGATTAATGATTAATTAATGATGCAACGCAAAAACCTTAACCTAC 956  
DB 127196 GCCCGCAAGCGGTGATGATGATTAATTAATGATGCAACGCAAAAACCTTAACCTAC 127137

QY 957 CCTTGAATGATGCAATTTTCTAAGAT-AGATTAGTCTTCGGGAACGCTTAACACAG 1015  
DB 127136 TCTTGACATCTTAAGAAAGCTCAAGATGATGATGATGCTTGTGCGGAATTAAGACAG 127077

QY 1016 TGTGCAATGCGCTGTGATGCTGATGCTGATGATGATGATGATGATGATGATGATGATG 1075  
DB 127076 TGTGCAATGCGCTGTGATGCTGATGCTGATGATGATGATGATGATGATGATGATGATG 127017

QY 1076 CAACCTGTCTAATTAATGCTATC-ATTTGTTGGCACTTAATGAGACTGCCGTGAC 1134  
DB 127016 CAACCTTAATCTTTGTTGGCACTTGTGCGGAATCAAGAGACTGCCAGTAT 126957

QY 1135 AAACCGAGAAAGTGGGATGATGCTCAAGTCTCTCATGAGCCTTAATGATGAGCTTCA 1194  
DB 126956 AAACCTGAGAAAGTGGGATGATGCTCAAGTCTCTCATGAGCCTTAATGATGAGCTTCA 126897

QY 1195 AGTAAATTAATGAGCGCTTAACAGAGGTTGCCAACCCCGAGGGAGCTTAATCTCAGAA 1254  
DB 126896 AGTGTCTTAATGAGCGCTTAACAGAGGTTGCCAACCCCGAGGGAGCTTAATCTCAGAA 126837

QY 1255 AGCGGTCTAGTCCGATCGAGATGCTGCACTGCACTCCGTAAGTCCGAATCGCTAGT 1314  
DB 126836 AGTACGTCTTAAGTCCGATTTGAGATGCTGCACTGCACTCCGTAAGTCCGAATCGCTAGT 126777

QY 1315 AATCGCGATCAGATGCTCGCGGTAAATACGTTCCCGGCTTTGTAACACACCGCCGTCA 1374  
DB 126776 AATCGCGATCAGAAATGTCGCGGTAAATACGTTCCCGGCTTTGTAACACACCGCCGTCA 126717

QY 1375 CACATGGAAGTGGTTTACCAAGAAAGCGATTAACCGTAAGAGAGGCGCTTGCA 1434  
DB 126716 CACATGGAAGTGGTTTGAACAGAAAGATGATTAACCTTTTGAAGGCGCTTACCA 126657

QY 1435 CGGTGATTAATGACTGAGGCTG 1457  
DB 126656 CGGTGATTAATGACTGAGGCTG 126634

```
RESULT 55
US-10-158-865-1
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Pleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40808)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44875)..(44875)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11924)..(11924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
```

```

NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13910)..(13910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

Query Match      66.7%; Score 971.8; DB 3; Length 1830121;
Best Local Similarity 81.3%; Pred.No. 0; Mismatches 267; Indels 7; Gaps 5;
Matches 1189; Conservative 0;

Db          1 ATTGAACGCTGGCGGCATCTTTACATGSCAAGTGAACGGCAGCACGG--ATGCTTGc 58
|||
770639 ATTGAACGCTGGCGGCAGCTTTAACATGCAAGTGAACGGACGATGAGAAGAAAGCTTCg 770698
|||||
59 ATC--TGATGGCGAGTGGCGGACGGGTGAGTAATGATCGGAACGTAATCCAGAAAGGGg 116
|||||
770699 TTTCTGTCGACGAGTGGCGGACGGGTGAGTAATGCTTGGGAATCTGCTTATGAGGGgg 770758
|||||
117 GGTAACGATCCAAAAGATGTGCTAATAACCGCATATACTCTAAGAGAGAAACGAGGGGATC 176
|||||
770759 GAATAACGACGGGAAACTGTGCTTAATAACCGCATATTATCGAAAGATGAAGATCGGGACT 770818

```

QY	177	GAAAGACCTTGCCGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTAAAGGC	236
Db	770819	GAGAGGCGCGATGCGCATATGAGATGAGCCCAAGTGGGATTAGTATGGTGGGTAAATGC	770878
QY	237	CTACCAAGCGACGATCGATGATTGGTGTGAGAGGACGACCGACACTGGGACTGAG	296
Db	770879	CTACCAAGCCTGCGATCTCTAGCTGGTCTGAGAGGATGACAGCCACACTGGAACTGAG	770938
QY	297	CACGCGCCGACACTCCTTACGGGAGCGAGAGTGGGGAATTTTGGACATGGCGAAGCCT	356
Db	770939	CACGCTCCAGACTCCTTACGGGAGCGAGAGTGGGGAATTTTGGCAATGGGGGAACCT	770998
QY	357	GATCCAGCAATGCGCGGTGAGTGAAGAAGCCTTCGGGTGTAAAGCTCTTCACTCGAG	416
Db	770999	GACCGAGCCATGCGCGTGAATGAAGAAGCCTTCGGGTGTAAAGTCTTTCGGTATGG	771058
QY	417	AAGAAAGGTATCGGTAAATATCGTGACTATGACGGTATCGAAGAAAGAACCCGGC	476
Db	771059	AGGAAGTGTGATGTATTATGATACATCAAAATTGACGTTAAATACAGAAAGAACCCGGC	771118
QY	477	TAACTACGCGCAGACGCGCGGTATACGTAGGGGTGCAAGCGTTATCGAATTAATCG	536
Db	771119	TAACTCCGTCGACGAGCGCGCGGTATATACGAGGGTTCGACGGTTATTCGAATATCG	771178
QY	537	GCGTAAAGGTGCGCAGCGCGCTTGTATGATGATGTGAATCCCGGGCTTAACTCG	596
Db	771179	GCGTAAAGGACACGACAGCGCGGTATTATTAAGTGAAGTGTGAAGCCCTGGGCTTAACTAG	771238
QY	597	GAAATGCGTTGAACCTACAGGCTTAGATGTGCGCAGAGGAGTGGCAATTCATGTGTA	656
Db	771239	GAAATGCAATTCAGACTGGGTAACTAGATTACTTTAAGGAGGGATGAATTTCCACTGTA	771298
QY	657	GCAGTGAATGGGTAGATATGAGAAACATCGATGGGGAAGCGCCTCTCGGTTAA	716
Db	771299	GCGGTGAATGGGTAGATATGAGAAATACCGAAGCGGAAGCGACCCCTTGGGAATG	771358
QY	717	CAGTACGCTCATGACGAAACGCTGGGGAGCAAAACGAGATTAGATACCTCGTAGTCCA	776
Db	771359	TACTGACGCTCATGTGCGAAACGCTGGGGAGCAAAACGAGATTAGATACCTCGTAGTCCA	771418
QY	777	CGCCCTAAACGATGTCACTAGTTGTTGGGCTTATTAGGCTTGGTAAAGAACTAACGC	836
Db	771419	CGCTGTAACGCTGTGCTGATTTGGGGATTGGGCTTAA-GACCTTGTGTGCGCGTAGCTAACGT	771477
QY	837	GTGAAGTTGACCGCGCTGGGAGATAGCGTGGCAAGATTAAACCTCAAAGGAATTGACGGG	896
Db	771478	GATTAATCAACCGCTCGGGAGATAGCGCGCAAGTTAACTCAATATTAATTGACGGG	771537
QY	897	ACCCGACAAAGCGGTGATTTATGTGGATTAAATTCATGTCAGCGCAAAAACCTTACTAC	956
Db	771538	GCCCGACAAAGCGGTGATGTGTTTAAATTCATGTCAGCGCAAAAACCTTACTAC	771597
QY	957	CCTTGCATGTAGCGAATTTCTAGAGAT-AGATTAGTCTTGGGAAAGCTAACACAGG	1015
Db	771598	TCTTGCATCTTAAGAAAGAGCTCAGAGATGAGGCTTGTGCTTCGGGAACTTAGAGACAGG	771657
QY	1016	TGCTGCATGGCTGTGCTGACGCTGTGTGCTGAGATTTTGGGTAAATCCCGCAACGACGG	1075
Db	771658	TGCTGCATGGCTGTGCTGACGCTGTGTGTAATTTGGGTAAATTCCTCGCAACGACGG	771717
QY	1076	CAACCTTGCATTAATTCGCAATC-ATTGGATTGGGCACTTTATATAGATGCGCGGTAC	1134
Db	771718	CAACCTTATCTCTTTTGTGCGACGACTTGTGCGGAACTCAAAAGAGATGCGCAGTGAT	771777
QY	1135	AAACCGAGAAAGTGGGATGACGTCAAGTCTCATGGCCCTTATGGGTAGGGCTTCAAC	1194
Db	771778	AAACTGGAAGAGTGGGATGACGTCAAGTCAATGACCCCTTACGAGTAGGGCTTACAC	771837
QY	1195	ACGTAAATACATGGCGCGTACAGAGGTTGGCCAAACCCGGAAGGGGAGCTAAATCAGAA	1254
Db	771838	ACCTGTCAATATGCGCTTAACAGAGGAAAGCGAAGTGTGGAAGTGAAGCAATCTCATTA	771897



QY	1255	AGCGGCTGTGTGTCCGGATCGGAGTCTTGCACTGCATCTCCGTGAAGTGTGGAATGGCTAGT	1314
Db	771898	AGTACGTCTTAAGTCCGGATTGGAGTCTGCACTGCATCTCAATGAAGTGGAAATGGCTAGT	771957
QY	1315	AATCGCGGATCAGCATGTCTGGGGTGAATACGTTCCGGGCTGTGTAACACACCGGCCGTCA	1374
Db	771958	AATGCGCATCAGAAATGTCTGGGGTGAATACGTTCCCGGGCCTTGTACACACCGCCGTCA	772017
QY	1375	CACCATGGAGTGGGTTTCACACAGACAGGTAGTCTAACCGTAAAGAGGGCGGCTTGCCA	1434
Db	772018	CACCATGGAGTGGGTTGTATACAGAAATGAATAGTTAACCTTTTGGAGGGCGTTTACCA	772077
QY	1435	CGGTGAGTTCACTAGCTGGGGTG	1457
Db	772078	CGGTATGATTTCACTAGCTGGGGTG	772100

```

RESULT 56
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2CID1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g

```

```

1  FEATURE:
2  NAME/KEY: misc_feature
3  LOCATION: (44416)..(44416)
4  OTHER INFORMATION: n equals a,t,c, or g
5  FEATURE:
6  NAME/KEY: misc_feature
7  LOCATION: (44905)..(44905)
8  OTHER INFORMATION: n equals a,t,c, or g
9  FEATURE:
10 NAME/KEY: misc_feature
11 LOCATION: (44975)..(44975)
12 OTHER INFORMATION: n equals a,t,c, or g
13 FEATURE:
14 NAME/KEY: misc_feature
15 LOCATION: (45593)..(45593)
16 OTHER INFORMATION: n equals a,t,c, or g
17 FEATURE:
18 NAME/KEY: misc_feature
19 LOCATION: (45732)..(45732)
20 OTHER INFORMATION: n equals a,t,c, or g
21 FEATURE:
22 NAME/KEY: misc_feature
23 LOCATION: (47036)..(47036)
24 OTHER INFORMATION: n equals a,t,c, or g
25 FEATURE:
26 NAME/KEY: misc_feature
27 LOCATION: (51334)..(51334)
28 OTHER INFORMATION: n equals a,t,c, or g
29 FEATURE:
30 NAME/KEY: misc_feature
31 LOCATION: (51602)..(51602)
32 OTHER INFORMATION: n equals a,t,c, or g
33 FEATURE:
34 NAME/KEY: misc_feature
35 LOCATION: (51805)..(51805)
36 OTHER INFORMATION: n equals a,t,c, or g
37 FEATURE:
38 NAME/KEY: misc_feature
39 LOCATION: (55369)..(55369)
40 OTHER INFORMATION: n equals a,t,c, or g
41 FEATURE:
42 NAME/KEY: misc_feature
43 LOCATION: (65313)..(65313)
44 OTHER INFORMATION: n equals a,t,c, or g
45 FEATURE:
46 NAME/KEY: misc_feature
47 LOCATION: (65309)..(65309)
48 OTHER INFORMATION: n equals a,t,c, or g
49 FEATURE:
50 NAME/KEY: misc_feature
51 LOCATION: (65313)..(65313)
52 OTHER INFORMATION: n equals a,t,c, or g
53 FEATURE:
54 NAME/KEY: misc_feature
55 LOCATION: (80024)..(80024)
56 OTHER INFORMATION: n equals a,t,c, or g
57 FEATURE:
58 NAME/KEY: misc_feature
59 LOCATION: (100091)..(100091)
60 OTHER INFORMATION: n equals a,t,c, or g
61 FEATURE:
62 NAME/KEY: misc_feature
63 LOCATION: (105121)..(105121)
64 OTHER INFORMATION: n equals a,t,c, or g
65 FEATURE:
66 NAME/KEY: misc_feature
67 LOCATION: (107248)..(107248)
68 OTHER INFORMATION: n equals a,t,c, or g
69 FEATURE:

```

NAME/KEY: misc\_feature  
LOCATION: (117136)..(117136)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (119750)..(119750)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (119924)..(119924)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (120038)..(120038)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (121344)..(121344)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (122167)..(122167)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (122336)..(122336)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (131340)..(131340)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (131360)..(131360)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (139910)..(139910)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (140398)..(140398)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (142750)..(142750)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (145058)..(145058)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (145171)..(145171)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (145942)..(145942)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (147197)..(147197)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (150841)..(150841)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (152500)..(152500)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature

Query Match 66.7%; Score 971 8; DB 3; Length 1830121;  
Best Local Similarity 81.3%; Pred. No. 0; Mismatches 267; Indels 7; Gaps 5;  
Matches 1189; Conservative

QY 1 ATTGAAGCTGGCGCATGCTTTTACACATGCAAGTCGAACGCGACACGG--ATGCTTGC 58  
DB 128095 ATTGAAGCTGGCGCATGCTTTTACACATGCAAGTCGAACGCGATCGAAGAAAGCTTGC 128036

QY 59 ATC--TGTGGCGAGTGGCGGACGCGGTGATGATGCAATCGAATCGATTCGAAGAGGG 116  
DB 128035 TTTCTTGTGACGATGGCGGACGCGGTGATGATGCTTGGAAATCTGCTTAATGAGGG 127976

QY 117 GGTAAAGCATGAAAGATGCTTAATCCGATATCTACTAAGAGGAAGAGGAGATC 176  
DB 127975 GATTAAGCATGAAAGATGCTTAATCCGATATCTACTAAGAGGAAGAGGAGATC 127916

QY 177 GAAAGACCTTGCGCTTTTGGAGCGCGCGATGCTGATTAAGCTAGTTGGTGGGTAAAGGC 236  
DB 127915 GAGAGCGCGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 127856

QY 237 CTACCAAGCGACGATCAGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 296  
DB 127855 CTACCAAGCGACGATCAGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 127796

QY 297 CACGGCCCAAGCTCTTACGAGGAGGACGATGAGGAAATTTTGAACAATGGGCGCAAGCCT 356  
DB 127795 CACGGTCCAGACTCTTACGAGGAGGACGATGAGGAAATTTTGAACAATGGGCGCAAGCCT 127736

QY 357 GATCCAGCAATGCGCGGAGAGGAAGGCGCTTGAGGTGTAAGCTCTTCACTGAG 416  
DB 127735 GACGAGCATGCGCGGAGAGGAAGGCGCTTGAGGTGTAAGCTCTTCACTGAG 127676

QY 417 AAGAAAGGTTACGTTAATATCTGATCTATGACGATGACGATGACGATGACGATGACGATG 476  
DB 127675 AAGAAAGGTTAATATCTGATCTATGATCATCAATATGATCAATATGATCAATATGATCA 127616

QY 477 TAACTACGTGCGACGACCGCGGTATATGATGATGATGATGATGATGATGATGATGATGATG 536  
DB 127615 TAACTACGTGCGACGACCGCGGTATATGATGATGATGATGATGATGATGATGATGATGATG 127556

QY 537 GCGTAAAGGTGCGGAGCGGCTTTGTAAGTGAATGGAATGCCCGGCTTAACTGG 596  
DB 127555 GCGTAAAGGTGCGGAGCGGCTTTGTAAGTGAATGGAATGCCCGGCTTAACTGG 127496

QY 597 GAATGCGTTGAACTCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 656  
DB 127495 GAATGCGTTGAACTCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 127436

QY 657 GCAAGTGAATGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 716  
DB 127435 GCGGTGAATGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 127376

QY 717 CACTGACGCTCATGACGGAAGCGTGGGAGGAACAGATTAAGATTAAGATTAAGATTAAGATTA 776  
DB 127375 TACTGACGCTCATGACGGAAGCGTGGGAGGAACAGATTAAGATTAAGATTAAGATTAAG 127316

QY 777 CGCCCTTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 836  
DB 127315 CGCCCTTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 127257

QY 837 GTGAAGTTGACCGCTGGGAGTACCGGTGCAAGATTAAGATTAAGATTAAGATTAAGATTA 896  
DB 127256 GATTAAGTTGACCGCTGGGAGTACCGGTGCAAGATTAAGATTAAGATTAAGATTAAG 127197

QY 897 ACCCGCAAGGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 956  
DB 127196 GCCCGCAAGGCGGTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 127137

QY 957 CCTTGACATGATGATTAATTTCTAGAGAT--AGATTAAGTCTTGGGAAAGCTTAACAGAG 1015  
DB 127136 TCTTGACATGATGATTAATTTCTAGAGAT--AGATTAAGTCTTGGGAAAGCTTAACAGAG 127077

Qy	1016	TGCTGCATGCGTGTGTGACGCTCGGTGTGAGATGTGGGTAAATCCCCGAAAGAGG	1075
Db	127076	TGCTGCATGCGCTGTGTGACGCTCGGTGTGAAATGTGGGTAAATGCCGCAACGAGCG	127017
Qy	1076	CAACCCCTGTCAATTAAATGGCATT-C-ATTGTGTTGGGCACTTTATATAGACTGCCGCTGAC	1134
Db	127016	CAACCCCTTATCTTCTTGTGTGCCAGCACTTGTGTGGGAACTCAAAAGAGACTGCCATGTAA	126957
Qy	1135	AAACCGAGAGAGGTGGGAGTACGTCAAGTCTCATGSCCTTATATGGGTAGGGCTTCA	1194
Db	126956	AAACTGGAGAGAGGTGGGATGACGTCAAGTCAATGAGCCCTTACGAGTATAGGGCTTAC	126897
Qy	1195	ACGTAAATCAATGGGCGCTTACAGAGGTTGCCAACCCGAGAGGGGGAGCTAATCTAGAA	1254
Db	126896	ACGTGTCTCAATGGGCTAATACAGAGGAGCCAGAGCTGTGAGAGTGAAGCAATCTATAA	126837
Qy	1255	AGCGCGTGTAGTCCGGATCGGAGTGTGCACACTCGACTCCGTGAATCGCTAGT	1314
Db	126836	AGTAGCTCTAAGTCCGGATTTGGAAGTGTGCACCTCGACTCAATGAATCGGAATCGCTAGT	126777
Qy	1315	AATCGCGATCAGCATGTCTCGCGTGAATACGTTCCCGGCTTTGTACACACCGCCCGTCA	1374
Db	126776	AATCCCGAATCAGAAATGTCTCGCGTGAATACGTTCCCGGCGCTTTGTACACACCGCCCGTCA	126717
Qy	1375	CACCATGGAGAGTGGGTTTACACAGAAAGCAGGTATTAACCGTAAAGAGGCGCTTGCCA	1434
Db	126716	CACCATGGAGAGTGGGTTTACACAGAAAGTAAATGCTTAACTTTTGGAGGGCGTTTACCA	126657
Qy	1435	CGGTGAGATTCAATGACTGGGGGTG	1457
Db	126656	CGGTATGATTCAATGACTGGGGGTG	126634

```

1 RESULT 57
2 US-08-239-810A-27
3 / Sequence 27, Application US/08299810A
4 / Patent No. 5721097
5 /
6 / GENERAL INFORMATION:
7 /
8 / APPLICANT: Rossau, Rudi
9 /
10 / APPLICANT: Van Heuverswyn, Hugo
11 /
12 / TITLE OF INVENTION: HYBRIDIZATION PROBES FOR THE
13 / TITLE OF INVENTION: DETECTION OF BRANHAMIELLA CATARRHALIS STRAINS
14 /
15 / NUMBER OF SEQUENCES: 28
16 /
17 / CORRESPONDENCE ADDRESS:
18 /
19 / ADDRESSEE: Merchant & Gould
20 /
21 / STREET: 3100 No. 5721097west Center
22 /
23 / CITY: Minneapolis
24 /
25 / STATE: MN
26 /
27 / COUNTRY: USA
28 /
29 / ZIP: 55402
30 /
31 / COMPUTER READABLE FORM:
32 /
33 / MEDIUM TYPE: Floppy disk
34 /
35 / COMPUTER: IBM PC compatible
36 /
37 / OPERATING SYSTEM: PC-DOS/MS-DOS
38 /
39 / SOFTWARE: Patentin Release #1.0, Version #1.25
40 /
41 / CURRENT APPLICATION DATA:
42 /
43 / APPLICATION NUMBER: US/08/299, 810A
44 /
45 / FILING DATE: 01-SEP-1994
46 /
47 / CLASSIFICATION: 435
48 /
49 / ATTORNEY/AGENT INFORMATION:
50 /
51 / NAME: Hillson, Randall A.
52 /
53 / REGISTRATION NUMBER: 31, 838
54 /
55 / REFERENCE/DOCKET NUMBER: 8076.70-US-WO
56 /
57 / TELECOMMUNICATION INFORMATION:
58 /
59 / TELEPHONE: 612-332-5300
60 /
61 / TELEFAX: 612-332-9081
62 /
63 / INFORMATION FOR SEQ ID NO: 27:
64 /
65 / SEQUENCE CHARACTERISTICS:
66 /
67 / LENGTH: 1485 base pairs
68 /
69 / TYPE: nucleic acid
70 /
71 / STRANDEDNESS: single
72 /
73 / TOPOLOGY: linear
74 /
75 / MOLECULE TYPE: DNA (genomic)

```

```

; ORIGINAL SOURCE:
; ORGANISM: Brannhamella catarrhalis
;
; IMMEDIATE SOURCE:
; CLONE: 16S rRNA Gene
;
US-08-299-810A-27

```

Query Match	66.2%	Score 965;	DB 2;	Length 1485;
Best Local Similarity	83.5%;	Pred. No. 0;		
Matches 1191; Conservative	0;	Mismatches 241;	Indels 12;	Gaps 77.

QY	17	ATGCTTTACATGCAAGTCGAAACGGCAGACCGGATCTTGGCATCTGGTGGCGAGTGGG	76
Db	1	AGGCTTAAACACATGGAAGTCGAAC-GAAGTTAGGAAGCTTGGTCTGATATTAGTGGCG	59
QY	77	GACGGGTAGTAAATCATCGAAGCGATTCAGAAAGGGGGGGTTAACGCATCGAAGATGT	136
Db	60	GACGGGTAGTAAATCTTAGAATCTGCTTAGTAGTGGGGGATTACTTGGGAAACCCAA	119
QY	137	GCTAATACCGCATATACTCTAAGAGAGAAAGCAGGGGATCGAAAGACTTGGCGCTTTGG	196
Db	120	GCTAATACCGCATACGACCTACGGGTGAAGGGGGGCTTT-----TAGCTCTGCTATTAG	174
QY	197	AGCGGCCAATGTCTTATTAAGTATGTTGGTGGGTAAAGGCTTAACCAAGGCGACGATCACT	256
Db	175	ATGAACCTTAAGTCGGATTAGCTGTGGTGGGTAAAGGCTTAACCAAGGCGACGATCTGT	234
QY	257	AGTTGCTGAGAGAGACGACCGACACTGGGACTGAGACACGGGCCAGACTCTCAACG	316
Db	235	AGCTGCTGAGAGAGATGATCAGCCACACTGGGACTGAGACACGGGCCAGACTCTCAACG	294
QY	317	GAGCGAGAGTGGGGAAATTTTGGACAATGGGCGCAAGCTGATCCAGCAATGCCCGCTGA	376
Db	295	GAGCGAGAGTGGGGAAATTTTGGACAATGGGCGCAAGCTGATCCAGCAATGCCCGCTGT	354
QY	377	GTGAAGAAAGGCTTCGGGTGTAAAGCTCTTCACTCCAGAAAGAAAGTTACGGTAAT	436
Db	355	GTGAAGAAAGGCTTTTGGTGTAAAGCACTTTTAAGTGGGAGAGAAAGCTTATGTGTAAT	414
QY	437	AATCGTAGCTCATGACGGTATCGACAGAAAGACACGGCTTAACCTACGTGCGACGACGG	496
Db	415	ACCCTAAGCCCTGACGTTATCCCAAGAAATAGCACCGGCTAACTGTGCGACGACGG	474
QY	497	CGGTAAATCGTAGGGTGCAGGCGTTAATCGAAATTACTGGGCGTAAAGGCTGCGACGGC	556
Db	475	CGGTAAATCAGAGGGGTGCAGGCGTTAATCGG-ATTACTGGCGTAAAGCGCGGTAGGT	533
QY	557	GCTTTGTAACTCAGATGTGAAATCCCCGGGCTTAACTTGGGAATTCGTTGAAATTACA	616
Db	534	GTTATTTAACTCAGATGTGAAAGCCCCGGGCTTAACTTGGGAATTCGATCTGATACCTGA	593
QY	617	AGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGATGACAGTGAATTCGTAAGAT	676
Db	594	TAACTAGAGTGTGAGAGGAGGAGTGAATTCAGGTGATGACGCTGTAATTCGTAAGAT	653
QY	677	ATGGAAGAACATCGATGGCGAAGGACGCTCTCGGCTTAACTGACGCTCATGACGA	736
Db	654	CTGAGGAATACCGATGGGAAAGGACGCTCCTGGCATCATCTGACACTGAGAGTGCGA	713
QY	737	AGCGTGGGAGCAACAGAGTTAGTACCCTGTGTGTCCAAGCCCTTAAACGATGTCACT	796
Db	714	AGCGTGGGTAGCAACAGAGTTAGTACCCTGTGTGTCCAAGCCCTTAAACGATGTCTACC	773
QY	797	AGTTGTGGGCTTAT-TAGGCTTGATTAAGAAAGCTAACGCGTGAAGTTGACCGGCTGGG	855
Db	774	AGTCTTGAGGTCTTTTAAAGACTTAGTGACGAGTTAAAGCAATATAGTGAACCGCTGGG	833
QY	856	GAGTACGGTCCGACAGATTAAACTCAAGGAATTGACGGGGACCCGCAACAGCGGTGAT	915
Db	834	GAGTACGGCCCGAAGTTAAACTCAATGATTTGACGGGGGCCGCAACAGGCGTGGAG	893
QY	916	TATGTGATTAATTGATGCAACCGGAAGAACTTACCTACCTTGACATGTAGCAAT	975
Db	894	CATGTGTTTATTTGATGCAACGGGAAGAACTTACCTGTGCTTGAACATGTAGCAATC	953

QY 976 TTTCAGATAGATTAGTG-CTTCGGGAAACGCTAACACAGGTGCTGATGCTGTCTCA 1034  
 Db 994 TTGAGAGATGAGAGATGCTTCCGGGAATTCACATACAGGTGCTGATGCTGTCTCA 1013  
 QY 1035 GCTGCTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCTTGTCAATTAATG 1094  
 Db 1014 GCTGCTGCTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCTTGTCTTAAGTAA 1073  
 QY 1095 CCATC-ATTGGGTGGGCACTTAATAGAGCTGCGGTGACAAACCGAGAGAGGTGGG 1153  
 Db 1074 CCAAGCTCTGCTGAGAACTTAAGATCTGCAAGTACAACTGAGAGAGAGCGGG 1133  
 QY 1154 ATGAGCTCAAGTCTCATAGGCTTATAGGTAGGGCTTACAGATTAATTAATGCGCT 1213  
 Db 1134 AGGAGCTCAAGTCAATGATGCTTATAGCAAGGGCTTACAGATTAATTAATGCTGCT 1193  
 QY 1214 ACAGAGGCTTGGCAACCCGAGAGGGAGCTTATCTCAGAAAGCGGTGTAGTCCGAT 1273  
 Db 1194 ACAAGGCTTGGCTACACAGCGATGATGCTATCTCAAAAAGCCATGTAGTCCGAT 1253  
 QY 1274 CGGATCTGCAACTGATCCGCTGAGAGTGGATGCTAGTATGCGGATCAGATGCTC 1333  
 Db 1254 TGGAGCTGCAACTGATCCGCTGAGAGTGGATGCTAGTATGCTAGTATGCAATGCT 1313  
 QY 1334 GCGGTGAATACGTTCCCGGCTTGTGTACACACCGCGCTCACACATGGGAGTGGTTTC 1393  
 Db 1314 GCGGTGAATACGTTCCCGGCTTGTGTACACACCGCGCTCACACATGGGAGTGGTTTC 1373  
 QY 1394 ACCAGAGAGGATGCTTAACCGTAAAGAGGCGCTTGCACGCTGATTAATTCATGCTGG 1453  
 Db 1374 ACCAGAGAGGATGCTTAAC--GCAAGAGGCGGATCACACGCTGGGTGATGCTGG 1431  
 QY 1454 GGTG 1457  
 Db 1432 GGTG 1435

RESULT 58  
 US-09-726-774-13  
 ; Sequence 13, Application US/09726774  
 ; Patent No. 6677153  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Iversen, Patrick L.  
 ; TITLE OF INVENTION: Antisense Antibacterial Method and  
 ; FILE OF INVENTION: Composition  
 ; FILE REFERENCE: 0450-0032.30  
 ; CURRENT APPLICATION NUMBER: US/09/726,774  
 ; CURRENT FILING DATE: 2000-11-29  
 ; PRIOR APPLICATION NUMBER: US 60/168,150  
 ; PRIOR FILING DATE: 1999-11-29  
 ; NUMBER OF SEQ ID NOS: 139  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13  
 ; LENGTH: 1487  
 ; TYPE: DNA  
 ; ORGANISM: Hemophilis influenza  
 ; FEATURES:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (1)...(1487)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-726-774-13

Query Match 66.1%; Score 963.8; DB 3; Length 1487;  
 Best Local Similarity 80.8%; Pred. No. 0;  
 Matches 1181; Conservative 0; Mismatches 273; Indels 7; Gaps 5;  
 QY 1 ATTGAAGCTGGGCGGATGCTTTACATGCAAGTCAAGCGGACGAG--ATGCTTGC 58  
 Db 28 ATTGAAGCTGGGCGGATGCTTTACATGCAAGTCAAGCGGATGCAAGCGGAGAAAGCTTGC 87  
 QY 59 ATC--TGGTGGCAGTGGGCGGACGGGTGATGATGCTGCAACGATTCAGAAAGAGGGG 116

Db 88 TTTCCTGTGACGATGGCGGACGGGTGATTAATGCTTGGGAATCTGCTTAATGAGGGG 147  
 QY 117 GGTACCGATGAAAGATGCTTAATCCGATTAATCTTAAGGAGAAAGCGGAGATC 176  
 Db 148 GATTAACGAGGAGAACTGCTGCTTAATACGGGTATTAATGGAAGTGAAGTGGCGGACT 207  
 QY 177 GAAAGACCTTGCGCTTTTGGAGCGGCGATGCTGATTAATGCTAGTGGTGGGTAAAGGC 236  
 Db 208 GAGAGCGGCAATGCCATAGATGAGTGAAGCCCAAGTGGGATTAAGTATGTTGGGTAAATGC 267  
 QY 237 CTACCAAGCGACGATCAGTATGTTGCTGAGAGGACCAACCACTGGGACTGAGA 296  
 Db 268 CTACCAAGCGGATCTCTAGCTGCTGAGAGGATGACCACTGGGACTGAGA 327  
 QY 297 CAGGCGCAAGTCTTACGAGGAGGACGATGGGGAATTTTGGAAATGGGGGCAAGCT 356  
 Db 328 CAGGCTCAAGTCTCTACGAGGAGGACGATGGGGAATTTTGGGAAATGGGGGCAAGCT 387  
 QY 357 GATCAGAGATGCGCGTGAAGTGAAGAGGCGCTTGGGTGTAAGCTCTTCACTGAG 416  
 Db 388 GACGAGCGATGCGCGTGAAGTGAAGAGGCGCTTGGGTGTAAGCTCTTCACTGAG 447  
 QY 417 AAGAAAGGTTACGTTAATATCTGATCTATGACGCTATGACAGAAAGACCGGC 476  
 Db 448 AGGAAAGGTTAATGTTAATATGACATCAATTAATGATTAATTAATGAGAAAGACCGGC 507  
 QY 477 TAACTACGTCAGAGCGCGGTAAATACGTAAGGTGCAAGGCTTAATGGAATTAATCG 536  
 Db 508 TAACTACGTCAGAGCGCGGTAAATACGTAAGGTGCAAGGCTTAATGGAATTAATCG 567  
 QY 537 GCGTAAAGGTGCGGAGCGGCTTGTAGTGAAGTGAATTCGCGGCTTAACTGG 596  
 Db 568 GCGTAAAGGTGCGGAGCGGCTTGTAGTGAAGTGAATTCGCGGCTTAACTGG 627  
 QY 597 GAATTCGCTTGAATTAATCAAGGCTAAGTGTGCAAGAGGAGTGAATTCATGTGTA 656  
 Db 628 GAATTCGCTTGAATTAATCAAGGCTAAGTGTGCAAGAGGAGTGAATTCATGTGTA 687  
 QY 657 GCAATGAATGCGTGAAGTATGAAAGAACATGATGGGAGGAGGAGGAGGAGGAGGAG 716  
 Db 688 GCGGTGAATGCGTGAAGTATGAAAGAACATGATGGGAGGAGGAGGAGGAGGAGGAG 747  
 QY 717 CACTACGCTATGACAGGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 776  
 Db 748 TACTACGCTATGACAGGAAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 807  
 QY 777 CGGCTTAACGATGCTCACTAGTATGTTGGGCTTATTAAGCTTGTGTAACGAACTAAC 836  
 Db 808 CGGCTTAACGCTGTGATTTGGGGTGGGTT-7AACCTGGGACCGGTAGCTAACGT 866  
 QY 837 GTGAAGTGAACCGCTGGGAGTACGCTGCAAGATTAATACTCAAGAAATTTGACGGG 896  
 Db 867 GATTAATGACCGCTGGGAGTACGCGCAAGGTTAAACTCAATGAATTTGACGGG 926  
 QY 897 ACCCGCAAGGCGGTGATTAATGATTAATTCATGCAAGCGGAAACCTTACTAC 956  
 Db 927 GCGMGCAAGGCGGTGATTAATGATTAATTCATGCAAGCGGAAACCTTACTAC 986  
 QY 957 CTTGACATGATGCGAATTTTCTAGAGT-AGATTAGCTTGGGAAAGCTTAACAGG 1015  
 Db 987 TCTTGACATCTTAAGAAAGCTCAGAGATGAGTGTGCTTGGGAACTTAAGAGAG 1046  
 QY 1016 TGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075  
 Db 1047 TGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106  
 QY 1076 CAACCTTGTCAATTAATTCGATC-ATTGCTTGGGCACTTAATGAGCTGCGGATG 1134  
 Db 1107 CAACCTTGTCAATTAATTCGATC-ATTGCTTGGGCACTTAATGAGCTGCGGATG 1166  
 QY 1135 AAACCGAAGAGGAGGAGTGAAGTCAAGTCTCAATGCGCTTAATGAGGAGGCTTCA 1194  
 Db 1167 AAACCGAAGAGGAGGAGTGAAGTCAAGTCTCAATGCGCTTAATGAGGAGGCTTCA 1226

QY 1195 ACGTAATACATGGCGCGGTACAGAGGTTGCCAACCCGCGAGGGGGAGCTTAATCTCAGAA 1254  
Db 1227 ACGGTCTAACATGGCGGTATACAGAGGAAAGCGAACTCGAGGTGAGCGAATCTCATPA 1286  
QY 1255 AGCGGTGTAGTCCGCGATCCGAGTCTGCAACTCGACTCCGTGAAGTGGAAATCGCTAGT 1314  
Db 1287 AGTACGTCTAGTCCGAGTTGGAGTCTGCAACTCGACTCGAATGAAAGTGGAAATCGCTAGT 1346  
QY 1315 AATGCGGATACAGCATGTTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCCTGCA 1374  
Db 1347 AATGCGGAATAGATGTTCGCGGTGAATACGTTCCCGGCTCTTGTACACACCGCCCTGCA 1406  
QY 1375 CACCATGGAGTGGGTTTCCACAGAAAGCAGTACTTAACCGTAAAGGAGGCGCTTGCA 1434  
Db 1407 CACCATGGAGTGGGTTTCCACAGAAAGTACTTAACCGTAAAGGAGGCGCTTGCA 1466  
QY 1435 CGGTGAGTTCATGACTGGG 1455  
Db 1467 CGGTATGATTCATGACTGGG 1487

RESULT 59  
US-10-278-942-1  
; Sequence 1, Application US/10278942  
; Patent No. 6720419  
; GENERAL INFORMATION:  
; APPLICANT: TAKARA BIO INC.  
; TITLE OF INVENTION: Sulfated fucan oligosaccharide  
; FILE REFERENCE: SAKA112  
; CURRENT APPLICATION NUMBER: US/10/278,942  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: JP 2001-325960  
; PRIOR FILING DATE: 2001-10-24  
; NUMBER OF SEQ ID NOS: 1  
; SEQ ID NO 1  
; LENGTH: 1506  
; TYPE: DNA  
; ORGANISM: Fucanobacter lyticus  
US-10-278-942-1

Query Match 66.1%; Score 963; DB 3; Length 1506;  
Best Local Similarity 81.3%; Pred. No. 0;  
Matches 1192; Conservative 0; Mismatches 265; Indels 10; Gaps 6;

QY 1 ATTGAACGCTGCGCGCATGCTTTACACATGCAATGCAACGCGACAGCATGCTTGAT 60  
Db 21 ATTGAACGCTGCGCGCATGCTTTACACATGCAATGCAATGCAATGCTTGAT 80  
QY 61 CT--GTTGGCGAGTGGCGGACGCGGTGAATGATGCAATGCAATGCAATGCAATGCAAT 118  
Db 81 ATTGGCGCTGACGCGCGGACGCGGTGAATGATGCAATGCAATGCAATGCAATGCAAT 140  
QY 119 TAAACGATCGAAAGATGTCTTAATACCGCATATCTTAAGGAGGAAAGCAGGCGATCGA 178  
Db 141 CAAACGTTGAAAGACATGCTTAATACCGCATATCTTAAGGAGGAAAGCAGGCGATCGA 200  
QY 179 AAG--ACCTTGCGCTTTTGGAGCGCGCATGCTGATTAAGCTTGGTGGGTTAAAGGC 236  
Db 201 TCGGAACCTTTGCGCATTTGATTAGCCCAAGTGAATTAAGCTTAAGGTTAAAGTATGGC 260  
QY 237 CTACCAAGGCGACATGATGATGTTGCTGAGAGGACGACACGACCACTGGGAATGAGA 296  
Db 261 TTACTTAGGCGACATGATGATGTTGCTGAGAGGATATCAAGCACTGGGAATGAGA 320  
QY 297 CACGCGCCAGACTCTTAACGAGGAGGACGAGTGGGAAATTTTGAACAATGGGCGCAAGCCT 356  
Db 321 CACGCGCCAGACTCTTAACGAGGAGGACGAGTGGGAAATTTTGAACAATGGGCGCAAGCCT 380  
QY 357 GATCCAGCAATGCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTCTTTCACTGAG 416  
Db 381 GATCCAGCAATGCGCGTGAAGTGAAGAGGCTTGGGTTGTAAGCTCTTTCACTGAG 440

QY 417 AAGAAAGTTACCGTAATATATCTGATCTCATGACGCTATCGACAGAAAGACACCGGC 476  
Db 441 AAGAAAGGTTGTATATATATCTGATCTCATGACGCTATCGACAGAAAGACACCGGC 500  
QY 477 TAACTAGTGCACAGACCGCGGTAAATCTGAGGTGCAAGCTTAAATCGAATTAATCTCG 536  
Db 501 TAACTAGTGCACAGACCGCGGTAAATCTGAGGTGCAAGCTTAAATCGAATTAATCTCG 560  
QY 537 GCGTAAAGGCGTGCACAGCGCGCTTGTAAAGTCAATGTAATGAAATCCCGGCTTAACTCG 596  
Db 561 GCGTAAAGGCGTGCACAGCGCGCTTGTAAAGTCAATGTAATGAAATCCCGGCTTAACTCG 620  
QY 597 GAATTCGCTTGAATCAACAGGCTAGAGTGTGCAAGGAGGTGAATTCATGTGTA 656  
Db 621 GAATTCGCTTGAATTCAGGCTAGAGTGTGCAAGGAGGTGAATTCATGTGTA 680  
QY 657 GCAGTAAATCGTGAAGATATGGAAGAACATGCAATGGCGAAGCAGCTTCTGGGTTAA 716  
Db 681 GCGGTGAATTCGTGAAGATATGGAAGAACATGCAATGGCGAAGCAGCTTCTGGACAGA 740  
QY 717 CACTGACGCTCATGACAGAAAGCGTGGGAGCAAAAGGATTAATACCTGTAGTCA 776  
Db 741 GACTGACCTGAGCAGAAAGCGTGGGAGCAAAAGGATTAATACCTGTAGTCA 800  
QY 777 CGCCTTAAACGATGTCAACTAGTTGT--TGGCCTTAATTAAGCTTGTGAACGAGCTAACG 835  
Db 801 CGCCTTAAACGATGTCAACTAGTTGT--TGGCCTTGTGAACGAGCTTGTGAACG 860  
QY 836 CGTGAAGTTGACCGCTGCGGAGTACGCTGCGAAGTAAATTAATCGAAGGATGACGG 895  
Db 861 CGTGAAGTTGACCGCTGCGGAGTACGCTGCGAAGTAAATTAATCGAAGGATGACGG 920  
QY 896 GACCGGACAAAGCGGTGATTAATGATTAATCGATGCAACGCGAAGGAACTTAACCTTA 955  
Db 921 GACCGGACAAAGCGGTGATTAATGATTAATCGATGCAACGCGAAGGAACTTAACCTTA 980  
QY 956 CCTTGAACATGATGCAATTTCTAGAGATGATTAATG--CTTGCGGAACGCTTAACACAG 1014  
Db 981 CCTTGAACATGATGCAATTTCTAGAGATGATTAATG--CTTGCGGTGAGGAGTATGACAG 1040  
QY 1015 GTGCTGATGCTGTGCTGACCTGCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGC 1074  
Db 1041 GTGCTGATGCTGTGCTGACCTGCTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGC 1100  
QY 1075 GCAACCTGTGCTTAATGTCATCATTTGGTGGGC--ACTTAATGAGACTGCGGCTG 1132  
Db 1101 GCAACCTGTGCTTAATGTCATCATTTGGTGGGC--ACTTAATGAGACTGCGGCTG 1160  
QY 1133 ACAACCGAGAGAGTGGGATGACGTCAAGTCTCATGCGCTTAATGAGTGGGCTTTC 1192  
Db 1161 ATTAACCGAGAGAGTGGGATGACGTCAAGTCTCATGCGCTTAATGAGTGGGCTTTC 1220  
QY 1193 ACACGTAAATACATGGCGCTGACAGAGGTTGCCAACCCGCGAGGGAGGCTTAATCTCAG 1252  
Db 1221 ACACGTCTACAAATGGACAGATGACAGAGGCTCAATCCGAGAGTGGAGGAAATCCAC 1280  
QY 1253 AAAGCGGTGCTATCCGGAATCGGAGTCTGCACTCGAATCCGGAAGTGGGAATTCGCTA 1312  
Db 1281 AAAGCGGTGCTATCCGGAATGAGTCTGCACTCGAATCCGGAAGTGGGAATTCGCTA 1340  
QY 1313 GTAATCGGATCAGCATGTGCGCGTGAATAGTCCCGGCTTGTGAACAACCGCCGCT 1372  
Db 1341 GTAATCGGATCAGCATGTGCGCGTGAATAGTCCCGGCTTGTGAACAACCGCCGCT 1400  
QY 1373 CACACATGGAGTGGGTTTCAACAAGAGGATGCTTAACGCT--AAGAGGCGCTT 1430  
Db 1401 CACACATGGAGTGGGTTTCAACAAGAGGATGCTTAACGCT--AAGAGGCGCTT 1460  
QY 1431 GCACGCTGAGATTCAGTCTGGGCTG 1457  
Db 1461 ACACCTTGTGATTCATGACTGGGCTG 1487

RESULT 60  
US-10-694-352-1  
Sequence 1, Application US/10694352  
Patient No. 6927289  
GENERAL INFORMATION:  
APPLICANT: TAKARA BIO INC.  
TITLE OF INVENTION: Sulfated fucan oligosaccharide  
FILE REFERENCE: SAKA112  
CURRENT APPLICATION NUMBER: US/10/694,352  
CURRENT FILING DATE: 2003-10-28  
PRIOR APPLICATION NUMBER: US/10/278,942  
PRIOR FILING DATE: 2002-10-23  
PRIOR APPLICATION NUMBER: JP 2001-325960  
PRIOR FILING DATE: 2001-10-24  
NUMBER OF SEQ ID NOS: 1  
SEQ ID NO: 1  
LENGTH: 1506  
TYPE: DNA  
ORGANISM: *Fucanobacter lyticus*  
US-10-694-352-1

Query Match 66.1%; Score 963; DB 3; Length 1506;  
Best Local Similarity 81.3%; Pred. No. 0;  
Matches 1192; Conservative 0; Mismatches 265; Indels 10; Gaps 6;

1 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAACGCGACGCGATGCTTGCAT 60  
21 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAACGCGACGCGATGCTTGCAT 80  
61 CT--GGTGGCGAGTGGCGGCGGAGTGAATGATCGGAACGTATCCAGAGAGGGGG 118  
81 ATTGGCGTGGAGCGGCGGCGGAGTGAATGCTTGGGAATATGCTTAATGGTGGGGGA 140  
119 TAAACCATGAAAGATGTCTTAATACCGCATATCTCTAAGAGAGAAACGAGGATTCGA 178  
141 CAACAGTTGMAACGATCTGTATACCGCATATGTCTACGAGCCAAAGAGGGGATTCCT 200  
179 AAG--ACCTTGGCTTTTGGAGCGCGGAGTGTCTGATGAGTGTGTGGGTAAAGGC 236  
201 TCGGAACCTTTCCCATTTGATTAAGCCAGATGAGTGAAGTGAAGTGAAGTGAAGTGC 260  
237 CTACCAAGCGGCGATCAATGATGTTGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 296  
261 TTACCTAGGCGAGATCTTACGCTGTTGAGAGAGATGATCAACCACTGGAGAGTGA 320  
297 CACGCGCCAGATCTCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 356  
321 CACGCGCCAGATCTCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 380  
357 GATCCAGCATATGCGGCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416  
381 GATGAGCCAGTGGCGGCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 440  
417 AAGAAAGGTTACGCTAATATATCTGATCATGACGCTATCGACAGAGAGAGAGAGAGAG 476  
441 AAGAAAGGTTACGCTAATATATCTGATCATGACGCTATCGACAGAGAGAGAGAGAGAG 500  
477 TAACTACGTCGAG 536  
501 TAACTACGTCGAG 560  
537 GCGTAAAGGTTGGCGGCGGCTTTGTAATGATGATGTAATGTAATGTAATGTAATGTAAT 596  
561 GCGTAAAGGTTGGCGGCGGCTTTGTAATGATGATGTAATGTAATGTAATGTAATGTAAT 620  
597 GAAATGCGTTGAAACTACAG 656  
621 GAACTGATTTTGAACCTGCGAATAATAGTTTGTAGAGGGTATGGAATTTTCAAGTGA 680  
657 GCACTGAATGCGTAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 716  
681 GCGGTGAATGCGTAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 740

717 CACTGACGCTCATGACAGAAAGCGTGGGAGCAAAAGAGATTAATGATACCTGTAGTCCA 776  
741 GACTGACACTGAGCAGCAAAAGCGTGGGAGCAAAAGAGATTAATGATACCTGTAGTCCA 800  
777 GCGCTAAACGATGCTCAACTAGTTGT--TGGGCTTAATTAAGCTTGGTAAACGACTAAG 835  
801 GCGCTAAACGATGCTCAACTAGTTGT--TGGGCTTAATTAAGCTTGGTAAACGACTAAG 860  
836 CGTGAAGTTGACCGCTGGGAGTACCGTGGCAAGATTAATAAATCTCAAAAGAAATTAACGG 895  
861 CGTGAAGTTGACCGCTGGGAGTACCGTGGCAAGATTAATAAATCTCAAAAGAAATTAACGG 920  
896 GACCGGCAACGCGTGGATTAATGGAATTAATGATGATGATGATGATGATGATGATGATG 955  
921 GCGCGGCAACGCGTGGATTAATGGAATTAATGATGATGATGATGATGATGATGATGATG 980  
956 CCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014  
981 CCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1040  
1015 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1074  
1041 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1100  
1075 GCAACCTTGTCTTAATTTGCCATCATTTGGTGGGC--ACCTTAATGATGATGATGATGATG 1132  
1101 GCAACCTTGTCTTAATTTGGCTGAGCGGATGATGATGATGATGATGATGATGATGATGATG 1160  
1133 ACAACCGGAGAGAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1192  
1161 ATAAACCGGAGAGAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1220  
1193 ACACGTTATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1252  
1221 ACACGTTATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1280  
1253 AAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312  
1281 AAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1340  
1313 GTAATCGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1372  
1341 GTAATCGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1400  
1373 CACACCATGAGAGTGGGTTTCAAG 1430  
1401 CACACCATGAGAGTGGGTTTCAAG 1460  
1431 GCGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1487  
1461 ACCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

RESULT 61  
US-09-726-774-1  
Sequence 1, Application US/09726774  
Patient No. 6677153  
GENERAL INFORMATION:  
APPLICANT: Iversen, Patrick L.  
TITLE OF INVENTION: Antisense Antibacterial Method and  
FILE REFERENCE: 0450-0032.30  
CURRENT APPLICATION NUMBER: US/09/726,774  
CURRENT FILING DATE: 2000-11-29  
PRIOR APPLICATION NUMBER: US 60/168,150  
PRIOR FILING DATE: 1999-11-29  
NUMBER OF SEQ ID NOS: 139  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 1  
LENGTH: 1450  
TYPE: DNA  
ORGANISM: *Escherichia coli*  
FEATURE:

NAME/KEY: misc\_feature  
LOCATION: (1)...(1450)  
OTHER INFORMATION: n = A,T,C or G  
US-09-726-774-1

Query Match: 64.8%; Score 944; DB 3; Length 1450;  
Best Local Similarity 81.3%; Pred. No. 0; Mismatches 258; Indels 8; Gaps 5;  
Matches 115; Conservative 0;

```

QY 1 ATTGAACGCTGGCGGCGATGCTTTACATGCAATGCAAGTCGACGCGACGAGATGCTTGA- 59
DB 19 ATTGAACGCTGGCGGCGGCGGCTTAACATGCAAGTCGACGCTTAACAGGAGCAGCTTGC 78
QY 60 ---TCTGTGCGGATGCGCGGCGGCTGATGATGCAATCGGAACCTATTCGAAAGGCG 116
DB 79 TCGCTTGTGCGGATGCGCGGCGGCTGATGATGCAATGCAAGCTTGAAGGCGG 138
QY 117 GGTACGCAATGCAAGATGCTATATACCGCATATCTCTAAGGAAAGGAGGCGATC 176
DB 139 GATTAATCTAGGAAACGCTAATATACCGCATATGCTCGCAAGCAAAAGGCGGAGC 198
QY 177 GAAAGACCTTGGCGCTTTTGAAGCGGCGGATGCTGATTAAGCTTGTGGGTAAAGGC 236
DB 199 TTAGGCGCTCTTGGCATGCGATGTCGCCAGATGGATTAAGTAAAGGTGGGTAAAGGC 258
QY 237 CTACCAAGGCGGATGCTGATGCTGCTGAGAGGACGACCGGCACTGCGGATCGAGA 296
DB 259 TCACCTAGGCGGATGCTGATGCTGCTGAGAGGATGACCGCAAACTGGAACCTGAGA 318
QY 297 CACGCGCGGATGCTGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 356
DB 319 CACGCTTCAAGATGCTGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 378
QY 357 GATCCAGCAATGCGCGCTGAGTGAAGAAAGGCGCTTGGGTGTAAGCTTTTCACTGAG 416
DB 379 GATGAGCGCATGCGCGCTGATGAAGAAAGGCGCTTGGGTGTAAGCTTTTCACTGAG 438
QY 417 AAGAAAGGTTACGCTTAATTAATCTGCTGATCATGACGCTTGAAGAAAGGAGCAGCGC 476
DB 439 AAGAAAGGTTACGCTTAATTAATCTGCTGATCATGACGCTTGAAGAAAGGAGCAGCGC 498
QY 477 TAACGAGGCGGAGGCGGCGGCTTAATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 536
DB 499 TAACGCGGCGGAGGCGGCGGCTTAATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 558
QY 537 GCGTAAAGGCTGCGGAGGCGGCTTGTGAAGTGAAGTGAATCCCGGCGCTTAACCTGG 596
DB 559 GCGTAAAGGCGGAGGCGGCGGCTTGTGAAGTGAAGTGAATCCCGGCGCTTAACCTGG 618
QY 597 GAATGCGCTTGAATCTACAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 656
DB 619 GAATGCGCTTGAATCTACAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 678
QY 657 GCAAGTGAATGCGTGAAGTATGAAGAAATCATGATGCGGAGGAGGAGGAGGAGGAGGAG 716
DB 679 GCGGTGAATGCGTGAAGTATGAGAAATCATGATGCGGAGGAGGAGGAGGAGGAGGAGGAG 738
QY 717 CACTGACGCTCATGACGAGAAAGCGTGGGAGCAAAACAGATTAAGTATGCTGATGCTCA 776
DB 739 GACTACGCTCATGACGAGAAAGCGTGGGAGCAAAACAGATTAAGTATGCTGATGCTCA 798
QY 777 GCGCTTAAACGATGCTCACTAGTGTGGGCTTATTAGGCTTGG- TAAAGAGCTAAGC 835
DB 799 GCGCTTAAACGATGCTCACTAGTGTGGGCTTATTAGGCTTGG- TAAAGAGCTAAGC 858
QY 856 CGTGAAGTGAACCGGCTGGGAGTACGCGTGAAGATTAAGTAAAGTGAAGTGAAGGAG 895
DB 859 CGTGAAGTGAACCGGCTGGGAGTACGCGTGAAGATTAAGTAAAGTGAAGTGAAGGAG 918
QY 896 GACCGGCAAGGCGGCTGATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTA 955
DB 919 G-GCGGCAAGGCGGCTGATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTA 977

```

```

QY 956 CCCTTGACATGATGAGAAATTTTGTAGAGT-AGATTAGTGTGCGGAGCGCTAACAG 1014
DB 978 GTCTTGACATGACGAGAAATTTTGTAGAGTATGAGATGCTGCTTGGAGACGTAGACAG 1037
QY 1015 GTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074
DB 1038 GTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1097
QY 1075 GCAACCTTGTCTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1133
DB 1098 GCAACCTTGTCTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1157
QY 1134 CAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1193
DB 1158 TAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1217
QY 1194 CACGCTTATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1253
DB 1218 CACGCTTATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1277
QY 1254 AAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1313
DB 1278 AAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1337
QY 1314 TAATGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1373
DB 1338 TAATGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1397
QY 1374 ACACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1414
DB 1398 ACACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1438

```

RESULT 62  
US-09-790-988-1  
Sequence 1, Application US/09790988  
Patent No. 6632935  
GENERAL INFORMATION:  
APPLICANT: SHIGENBU, SHUJI  
APPLICANT: WATANABE, HIDEKI  
APPLICANT: HATTORI, MASAHIRA  
APPLICANT: SAKAKI, YOSHIYUKI  
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
FILE REFERENCE: 081356/0159  
CURRENT APPLICATION NUMBER: US/09/790,988  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: JP2000-107160  
PRIOR FILING DATE: 2000-04-07  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 640681  
TYPE: DNA  
ORGANISM: Buchnera sp.  
US-09-790-988-1

Query Match: 64.7%; Score 942.8; DB 3; Length 640681;  
Best Local Similarity 79.9%; Pred. No. 3.2e-313;  
Matches 1175; Conservative 0; Mismatches 282; Indels 13; Gaps 5;

```

QY 1 ATTGAACGCTGGCGGCGGCTTTTACATGCAATGCAAGTCGACGCGACGAGG-----TGC 54
DB 274064 ATTGAACGCTGGCGGCGGCTTTTACATGCAATGCAAGTCGACGCGACGAGGAGGAGG 274123
QY 55 TTGCATCTGTGCGGAGTGGCGGAGCGGCTGATGATGCTGCAATCGTAATCGTAAGAGG 114
DB 274124 CTCTCTTTTGGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 274183
QY 115 GGGGTAACGATGCAAAAGATGCTTAATCCGCTATATCTTAAGAGGAGGAGGAGGAGGAG 174
DB 274184 GGGGTAACGATGCAAAAGATGCTTAATCCGCTATATCTTAAGAGGAGGAGGAGGAGGAG 274243
QY 175 TCGAAAGACCTTGGCGCTTTTGAAGCGGCGGATGCTGATTAAGTGAAGTGGGTAAAG 234

```



Db	274244	CTTTTGGCCTCATCTCTTTTGGATGAAACCAGACGAGATTAGCTTGTGGTAGAGTAA	274303
OY	225	GCTTACCAAGGCGACGATCAGTAGTGTGTCTGAGAGGACGACCGACCACTGGGACTGA	294
Db	274304	GCTTACCAAGGCGACGATCTTAGCTGGTCTGAGAGATTAACGACCACTGGAACTGA	274363
OY	295	GACACGGCCAGACTTCTACGGGAGGACAGTGGGGAAATTTTGGACATYGGGCCAAGC	354
Db	274364	GACACGGTCCAGACTTCTACGGGAGGACAGTGGGGAAATTTTGGACATYGGGCCAAGC	274423
OY	355	CTGATTCACGCAATGCCCGGTAGTGAAGAAAGCCCTTCGGGTTGTAAAGCTCTTTCAGTGG	414
Db	274424	CTGATGACAGCTATGCGCGCTGTATGAAGAAAGCCCTTAGGGTTGTAAAGTACTTTCAGCGG	274483
OY	415	AGAAAGAAAAGTTACGGTAAATATATCGTACCTATGACCGGTATCGACAGAAAGACCGG	474
Db	274484	AGAGAGAAAATTAATACTAATATTTTATTTTCGTGACCTTACCGCAGAAAGACCGG	274543
OY	475	GCTAATCAGTCCAGACAGCCGCGGTAAATACGTAGGGTCCAAAGCTTAATTCGAATTACT	534
Db	274544	GCTAATCTCGTCCAGACAGCCGCGGTAAATACGGAGGGTCCAAAGCTTAATCAGAAATTACT	274603
OY	535	GGGCGTAAAGGCTGCGCAGGCGGCTTTGTAAATGTGAAATATCCCGGGCTTAACCT	594
Db	274604	GGGCGTAAAGACCGCGCTAGTGGTTTTTAAAGTCAGAGTGTGAATCCCTAAGCTCAACCT	274663
OY	595	GGGAATTTGCGTTTGAACATCAAGAGCTAGAGTGTGCAAGGAGGCTGGAAATTCATGTG	654
Db	274664	AGGAATCTGATTTGAACCTGGAACCTAGAGTTTGTAAAGGAGGTGAATTTCTAGGTG	274723
OY	655	TAGCAGTGAATGCGTAGAGATATGGAAGAACATCGATGGCGAAGCGACTTCTGGGTT	714
Db	274724	TAGCGGTGAATGCGTAGATATCTGGAAGAAATCCGTGGCGAAACCGGCTCTAAACG	274783
OY	715	AAACCTGACGCTCATGCAAGAAAGCTGGGGAGCAACAGAGTTAATACCTCTGTAAGTC	774
Db	274784	AAAACCTGACGAGGCGCGAAGACGTGGGGACCAACAGAGTTAATACCTCTGTAAGTC	274843
OY	775	CACGCCCTTAAGATGTCAACTGATGTTGGGGCTTATTAGGCTTG- GTAAACAACCTAA	833
Db	274844	CATGCGGTAAACGATGTGCACTTGAAGGTTGTTTCCAGAGAAAGTACCTCCBAAGCTAA	274903
OY	834	CGCGTGAAGTTGACCGCTGGGGAGTAGACGCTCGCAAGATTAAACTCAAGGAATTGACG	893
Db	274904	CGCATTAAGTCCAGCCCGCTGGGGAGTAGACGCGCCCAAGGCTTAATAATGAATTGACG	274963
OY	894	GGGACCCCGACAAAGCGGTGATTATGTGATTATTCGATGCAACGCGAAAAACCTTACC	953
Db	274964	GGGGCCCGACAAAGCGGTGAGCATGTGTTTAATTCGATGCAACGCGAAAAACCTTACC	275023
OY	954	TACCTTTGACATGTAGCGAAATTTTCTAGAGATAGATTAGTG- CTTCGGGAAACGCTAAC	1012
Db	275024	TGCTCTTTGACATGCAAGAAATTTCTTAGAAATTAAGAAAGTGCTTCGGAGCTGTGAGAC	275083
OY	1013	AGGTGCTGACATGGCTGTGCTGAGCTGTGCTGAGAGATGTGGGTTAAGTCCGCAACGA	1072
Db	275084	AGGTGCTGACATGGCTGTGCTGAGCTGTGCTGAGAGATGTGGGTTAAGTCCGCAACGA	275143
OY	1073	GGCGAACCCCTTGTCAATTAATTTGCCATC- AATTGGTTGGGCACTTTAATGAGATGCCGT	1131
Db	275144	GGCGAACCCCTTATCCCTGTTTGGCAGAGGGTTCCGCGGAACTCAGAGAGATCGCCGT	275203
OY	1132	GACAAACCGAGGAAAGGTGGGAGTAGACTCAAGTCTCATAGGCCCTTATGGTATGGGCTT	1191
Db	275204	TATTAACCCGAGGAAAGGTGGGAGTAGAGTCAAGTCAATGATCATAGGCCCTTATGACCGGCTA	275263
OY	1192	CACACGTAATACATGGCGGTAACAGAGGGTTGCAACCGCGAGGGGAGAGTATCTCA	1251
Db	275264	CACACGCTACATGGTTTATATCAAGAGAGCAAAATCTGCAAAAGACAGCAAACTCTCA	275323
OY	1252	GAAAGCGCTGTAAGTCGCGATCGAGTCTGCACTGCACTCGGTGAAGTCGAATCGCT	1311

Dd	275324	TAAAGTAATGCTACTCCGGACGTGAAGTCTGCAGACTCGACCTCACAAGAAATCGCT	275383
Oy	1312	AGTATCCGGATTCAGCATGTTCGGGTGAATACGTTCCCGGCTTTGTACACACGCCGCCG	1371
Dd	275384	AGTATCGTGATTCGAATGCCACGGTGAATACGTTCCCGGCTTTGTACACACGCCGCCG	275443
Oy	1372	TCACACCATTGGAGTGGGTTTTCACCAAGAAGAGTAAAGTAAACCGT-----AAGAGGGCG	1427
Dd	275444	TCACACCATTGGAGTGGGTTTGAAGAAAGAACAGTAACTCTTAACCTTTTAAAGAGAGCGC	275503
Oy	1428	CTTGCCACGGTGCAGATTCATGACTGGGGTG	1457
Dd	275504	CTTACCACCTTTGTGATTCATGACTGGGGTG	275533
RESULT 63			
US-10-761-509-1			
Sequence 1, Application US/10761509			
Patent No. 6919197			
GENERAL INFORMATION:			
APPLICANT: Gerber, John F			
APPLICANT: White, James H			
TITLE OF INVENTION: Materials and Methods For The Efficient Production Of Pasteuria			
FILE REFERENCE: GER-100XC1			
CURRENT APPLICATION NUMBER: US/10/761,509			
CURRENT FILING DATE: 2004-01-20			
PRIOR APPLICATION NUMBER: US/09/636,185			
PRIOR FILING DATE: 2000-08-10			
PRIOR APPLICATION NUMBER: US 60/148,154			
PRIOR FILING DATE: 1999-08-10			
NUMBER OF SEQ ID NOS: 1			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO 1			
LENGTH: 1455			
TYPE: DNA			
ORGANISM: Unknown			
FEATURE:			
OTHER INFORMATION: Helper Factor DNA is not Pasteuria but has 98% homology with			
OTHER INFORMATION: Enterobacter cloacae and Pantoea asp. and others.			
FEATURE:			
NAME//KEY: misc_feature			
LOCATION: (99)..(99)			
OTHER INFORMATION: n = a, c, g, or t.			
FEATURE:			
NAME//KEY: misc_feature			
LOCATION: (102)..(102)			
OTHER INFORMATION: n = a, c, g, or t.			
US-10-761-509-1			
Query Match                  63.5%; Score 924.8; DB 3; Length 1455;			
Best Local Similarity      82.8%; Pred. No. 1,1e+08;			
Matches 1091; Conservative    0; Mismatches 224; Indels     3; Gaps     3;			
Oy	60	TCCTGTGGCGAGTGGCGGACGGGTGAGTAATGATCGGAACTATCCAGAAAGGGGGCT	119
Dd	50	TTTGGCGGCGAGCGGCGGACGGGTGAGTAATGCTGGGAACTACTGANGANTGGGAT	109
Oy	120	AACGATCCAAAAGATGTCTTAATCCGCCATATCTCTAAGGAGAAAGCAGGGATCGAA	179
Dd	110	CCTACTGGAACAGTTGCTTAATACGCATTAACCTTCMAAGACCAAAGAGGGGACTTC	169
Oy	180	AGACTTGGCGCTTTTGGAGCGGCGGATGTCTGATTAGCTAGTTAGTGGGGTAAAGGCTTA	239
Dd	170	GGGCTCTTGGCCATCAGATGTGCCCCAGATGGGATTAGTAGTGGGGTAAAGGCTCA	229
Oy	240	CCAAGCGACGATCAGTATGTTGCTTGAGAGACACACAGCCACACATGGGCTAGACAC	299
Dd	230	CTTAGCGACGATCCCTAGCTGTCTGAGAGATGACACAGCACACTGAACTGAGACAC	289
Oy	300	GGCCCGACGCTCTTACGGGAGGACAGTGGGGAAATTTGGCAATGGGCGCAAGCTGAT	359
Dd	290	GGTCCAGACTCTTACGGGAGGACAGTGGGGAAATTTGGCAATGGGCGCAAGCTGAT	349

QY 360 CCAGCATGCGCGCTGAGTGAAGAGGCGCTTCGGGTTGAAGCTCTTTCAGTCGAGAAG 419  
DB 350 GCAGCCATGCGCGCTGATGAAGAGGCGCTTCGGGTTGAAGCTCTTTCAGCGGAGG 409  
QY 420 AAAAGTTACGGTTAAATATCTGATCTGATGCGGTTATGACAGAGAGACCGGCTAA 479  
DB 410 AAGGCTGTAAGGTTAATACTCAGCGATTGACCTTACCCGAGAGAGACCGGCTAA 469  
QY 480 CTACGTGCGAGAGCGCGGTTAATAGTGGTGCAGAGGTTAATCGGAATTAATCTGGCG 539  
DB 470 CTCCTGTCAGAGCGCGGTTAATAGGAGGTGCAAGGTTAATCGGAATTAATCTGGCG 529  
QY 540 TAAAGGTCGAGCGCGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATG 599  
DB 530 TAAAGCGACGAGCGCGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATG 589  
QY 600 TTGCGTTTGAATCTAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 659  
DB 590 CTGATTCGAAACTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 649  
QY 660 GTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 719  
DB 650 GTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 709  
QY 720 TGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779  
DB 710 TGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 769  
QY 780 CCTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838  
DB 770 CGTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 829  
QY 839 GAACTTGAACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 898  
DB 830 TAAAGTCAGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 889  
QY 899 CCGCAGAGCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958  
DB 890 CCGCAGAGCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 949  
QY 959 TTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017  
DB 950 TTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1009  
QY 1018 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1077  
DB 1010 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1069  
QY 1078 ACCCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1136  
DB 1070 ACCCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1129  
QY 1137 ACCGAGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1196  
DB 1130 ACTGAGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1189  
QY 1197 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1256  
DB 1190 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1249  
QY 1257 GCGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1316  
DB 1250 TCGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1309  
QY 1317 TCGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1374  
DB 1310 TCGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1367

RESULT 64  
US-08-114-695A-2  
; Sequence 2, Application us/08114695A  
; Patent No. 5508193

GENERAL INFORMATION:  
APPLICANT: Mandelbaum, Raphael T.  
APPLICANT: Mactech, Lawrence P.  
TITLE OF INVENTION: DEGRADATION OF S-TRIAZINES IN SOIL AND  
TITLE OF INVENTION: WATER  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCHWEMAN, LUNDBERG & MOESSNER, P.A.  
STREET: 3500 IDS CENTER  
CITY: MINNEAPOLIS  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/114,695A  
FILING DATE: 31-AUG-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MUEHLING, ANN M.  
REGISTRATION NUMBER: 33,977  
REFERENCE/DOCKET NUMBER: 600.268US1  
TELEPHONE: 612-339-0331  
TELEFAX: 612-339-3061  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1473 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: rRNA  
ORIGINAL SOURCE:  
ORGANISM: Pseudomonas sp. Atrazine-Degrading Isolate  
US-08-114-695A-2  
Query Match 60.1%; Score 876.2; DB 2; Length 1473;  
Best Local Similarity 64.1%; Pred. No. 7.3e-292;  
Matches 921; Conservative 225; Mismatches 277; Indels 14; Gaps 9;  
QY 17 ATGCTTTACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 76  
DB 42 AGGCTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 100  
QY 77 GAGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 135  
DB 101 GAUGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 160  
QY 136 TGTATATACCGCATATATCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 195  
DB 161 CGUAAUACCGCAUACGUCUCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 220  
QY 196 GAGCGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 255  
DB 221 SAUAGCCUAGGCGGCUUAGCUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 280  
QY 256 TAGTTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 315  
DB 281 UAAUCUGUCUGAG 340  
QY 316 GAGGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 375  
DB 341 GAGGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 400  
QY 376 AGTGAAGAGCGCTTCGGGTTGTAAAGCTTTTCAATGTCGAGAGAGAGAGAGAG 435  
DB 401 UGUAGAGAGAGGUCUUCGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 460

436 TAATCGTGA CTCATGACGGTATCGACAGAAAGCACCGGCTTAATCTAGCTGCCAGACGCC 495  
461 UACCUUGCUNUUUUGACGUAUCCAAAGAAUAAAGCACCGGCUAACUUCUGGACAGACGCC 520  
496 GCGGTAATTC-GTAGGGTCCAAACGCTTAATCCGAATTAATCTAGGCGGTAAGGGTGGCGACG 554  
521 GCGGTAUUAUCUAAAGGUCUAGCGUUAUUCGAAUUAUACUGAGGAGGAAAGCGCGUAGG 580  
555 CGGCTTTGAATGACGATGTAATAATCCCGGCTTAATCTAGGGAATTTGGCTTTGAAACTA 614  
581 UGGUUGUAUAGUAUUGAUAUCCCGGCGUACAACUUGGGAGACUAGUACAAUACAU 640  
615 CAAGCTAGAGTGTGACAGAGGAGGTGGAATTCATGTGTACAGTGAATGCTGTAAG 674  
641 CCUAGCUAAGGACGCGUAGAGGUGUGGAUUAUUCUGUUAAGCGGUAUAGCGUAGAU 700  
675 ATATGGAAGAATCATGAGCCGAAGGCAACCTCTGCGGTTAACATGACGCTCATGACAG 734  
701 AUAAGAAAGAAACACAGUGGCGGAAAGGCGACACCTUAGACUAGACUAGCCUAGGUGCG 760  
735 AAAGCGTGGGAGCAACAGATTAATGATACCTGTGATGACAGCCCTAAACGATGTCAA 794  
761 CAAGC-UAGGAGCTMAACAGGAUUAUACCTUGUAUUCACGCGGUCACAGAUUCCGA 819  
795 CTAGTTGTGGGCTTATTAGGCTTGATTAAGGCTTAACGAGCTAACGCGTGAACGCTTGG 854  
820 CUAGCGGUUGGGAUCCUAGAUUUGGCGCAGUAAGGCGAUAAGUCGACCGCCTGG 879  
855 GAGATACGCTCGCAATTAATACTCAAGAAATTAAGCGGGGACCGGCAACAGCGGTGA 914  
880 GCGGUUACGCGCCGCAAGGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 936  
915 TTATGAGATTAATTAATGATGACAGCGGAAACCTTAACCTTGAACATGATGAGGAAT 974  
937 GCAGUGGUAUUAUUGAANNAACGGAAGAACCUAACCTUGGCTUAGCAUUGCCGAU 996  
975 TTTCTAGAG-ATAGATTAGTCTTCGGAACGCTAACACAGGTCTGATGCTGTCTC 1033  
997 CTGACAGAGCAUCCGAGAGTCTUUCGGAUUCGGAACACAGUGUGUAGUUGGUGUCC 1056  
1034 AG--CTGCTGTGATGATGTTGGGTTAAGTCCCGCAACGACGCGCAACCTTGTCTTA 1090  
1057 AGGUAUGUGUGUGUGAUGUUGGUAUUGUCCCGTAACGACGCGCAACCTUUGUCCUA 1116  
1091 ATTGCCATCA--TTTGGTTGGGCACTTAATGAGACTGCGCGGTGACAAACCGGA-GGAG 1147  
1117 GUUACCAAGCAACGUAUGUGGAGGACUCUAAGAGACUUGCCGUGACAAACCGAGCGAG 1176  
1148 GTGGGATGACGTCAGTCTCTCAATGAGCTTATGAGGCTTCAACGCTTAATTAATG 1207  
1177 GUGGGAUAGACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1236  
1208 GCGGTAACAGAGGTTGCCAACCCGCAAGGGGAGCTTAATCTCAAGAAAGCGCTGTAAT 1267  
1237 GUGCGUAACAGAGGUGUCCCAAGCGCGAGGUGAGCTUAUCCCAAGAAACCGAGGAGU 1296  
1268 CCGATATGAGATCTGCAATCTCGTGAAGTGGAAATCGCTAATCGCGGATCAG 1327  
1297 CCGGAUUGCAAGUCUGCAACUAGCUAGUAGUGCGGAUUCGUAAGUAUUCUGUAUACAG 1356  
1328 CATTCGCGGTGAATTAATGTTCCCGGCTTGTGTAACAACCGCCGCTCAACCATGGAAGT 1387  
1357 AAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 1416  
1388 GGTTCACCAAGACAGGATGCTTAACCTGAAGAGAGGCGCTTCCACGCTGAGATT 1444  
1417 GGUGGUCUACAGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1473

RESULT 65  
US-10-808-807-16  
; Sequence 16, Application US/10808807  
; Patent No. 6929928

GENERAL INFORMATION:  
; APPLICANT: E. I. du Pont de Nemours and Co., Inc.  
; APPLICANT: Cheng, Qiong  
; APPLICANT: Tao, Luan  
; APPLICANT: Sedkova, Natalia  
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS  
; FILE REFERENCE: CL2365 US NA  
; CURRENT APPLICATION NUMBER: US/10/808,807  
; PRIOR FILING DATE: 2004-03-24  
; PRIOR APPLICATION NUMBER: US 60/477,874  
; PRIOR FILING DATE: 2003-06-12  
; PRIOR APPLICATION NUMBER: US 60/527,083  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 16  
; LENGTH: 1324  
; TYPE: DNA  
; ORGANISM: Pantoea agglomerans strain DC404  
US-10-808-807-16  
Query Match 60.1%; Score 875.4; DB 3; Length 1324;  
Best Local Similarity 81.9%; Pred. No. 1.3e-291;  
Matches 1082; Conservative 0; Mismatches 231; Indels 8; Gaps 6;  
6 ACCTGCGCGCATGCTTTAACAATGCAATGCAAGCGGAT--GCTTGATCT- 62  
4 ACCTGCGCGCATGCTTTAACAATGCAATGCAAGCGGAT--GCTTGATCT 63  
63 -GGTGGAGTGGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121  
64 CGCTGACAGTGGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 123  
122 CCATGGAAGATGCTTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 181  
124 CTAAGTGAAGATGCTTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 183  
182 ACCCTGCGCTTTGAGACCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 241  
184 GCTCTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243  
242 AAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301  
244 TAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 303  
302 CCCAGATCTCTACGAGGAGCGAGTGGGAAATTTGAAATGCGCGCAAGCTGATTC 361  
304 TCAGATCTCTACGAGGAGCGAGTGGGAAATTTGAAATGCGCGCAAGCTGATTC 363  
362 AGCAATGCGCGTGAATGAAAGGCTTCCGCTTGAAGCTCTTCACTGAGAGAA 421  
364 AGCAATGCGCGTGAATGAAAGGCTTCCGCTTGAAGCTCTTCACTGAGAGAA 423  
422 AAGTTAAGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481  
424 GGTGTAAGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483  
482 ACCTGCAAGAGCGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541  
484 CGGTGCAAGAGCGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543  
542 AAGGATGCGAGCGCGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601  
544 AAGGATGCGAGCGCGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 603  
602 GCTTTGAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661  
604 GCAATGGAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663  
662 GAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721  
664 GAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723

QY	722	ACGCTCATGCACGAAAGCGTGGGAGACAAACAGAGATTAAATACCTCGTGTAGTCCAGGCC	781
Db	724	ACGCTCAGGTGGGAAAGCGTGGGAGACAAACAGAGATTAAATACCTCGTGTAGTCCAGGCC	783
QY	782	TAAAGCATGTCAACCTAGTGTGGGCTTATTAATGAGCTTGG- TAAAGAGCTAACGCGTGA	840
Db	784	TAAAGCATGTCCAGCTTGGAGGTTGTTCCTCTGAGAGATGGCTTCGGAGCTAACGCGTTA	843
QY	841	AGTTGACCGCCTTGGGAGTACCGGTCCGCAAGATTAAACCTCAAGAAATTGACGGGGACCC	900
Db	844	AGTCGACCGCCTTGGGAGTACCGGCCCGCAAGGTTAAACCTCAATGAAATTGACGGGGACCC	903
QY	901	GCAACAAGCGGTGATTAATGTGATTAATTCGATGCACCGGAAAAACCTTAACCTACCTT	960
Db	904	GCACAAGCGGTGAGCATGTGTTTAATTCGATGCACCGGAAACCTTAACCTGACCTT	963
QY	961	GACATGTAGCGAATTTCTAGAGATGATTAATG-CTTGGGGAACGCTAACAGGAGCT	101
Db	964	GACATTCACGGAATTCGCAAGATATCCCTTAGTGCTTGGGGAACGCTGAGACAGGTGCT	102
QY	1020	GCATGGCTGTGCTGCAGCTCGTGTCTGAGATGTTGGGTTAAAGTCCGCAACGAGCGCAAC	107
Db	1024	GCATGGCTGTGCTGCAGCTGTGTGTGAATGTGGGTTAAAGTCCGCAACGAGCGCAAC	108
QY	1080	CCTTGTCAATTAATGTCATC-ATTGGTGGGCACTTTAATGAGACTGCCGCTGACAAAC	113
Db	1084	CCTTATCTTGTGTGTCAGCGAATCCGTCGGGAACCTCAAGGAGACTGCCGCTGATTAAC	114
QY	1139	CGGAGGAAGCT-GGGAGTACGCTCAAGTCTCAATGAGCCCTTAATGGGTAGGGCTTACAG	119
Db	1144	CGGAGGAAGGTGGGGATGACGTCGAATCAATGAGCCCTTAACGCGCAAGGCTACACAG	120
QY	1198	TAAATCAATGGGCGCTACAGAGGGTTCGCAACCCGCGAGGGGAGCTTAATCTCAGAAAGC	125
Db	1204	TGCTACAAATGGGCTATACAAAGAGAGCGAAGCTCGGAGAGCAAGGAGCACTCATTAAGT	126
QY	1258	GCGTGTAGTCCGAGTCGAGTCTGCACTCGACTCCGTGAAGTCGGAATCGTTAGTAAT	131
Db	1264	ACGTGTAGTCCGAGTCGAGTCTGCAACTCGACTCCGTGAAGTCGGAATCGTTAGTAAT	132
QY	1318	C 1318	
Db	1324	C 1324	

RESULT 66  
US-09-347-001-1  
; Sequence 1, Application US/09347001A  
; Patent No. 635177  
; GENERAL INFORMATION:  
; APPLICANT: MIHARA, Yasuhiro  
; APPLICANT: TAKEUCHI, Sonoko  
; APPLICANT: JOJIMA, Yasuko  
; APPLICANT: TONOUCHI, Naoto  
; APPLICANT: FUDOU, Ryoosuke  
; APPLICANT: YOKOBEKI, Kenzo  
; TITLE OF INVENTION: NOVEL MICROORGANISM AND METHOD FOR PRODUCING XYLITOL OR  
; FILE REFERENCE: 0010-1015-0  
; CURRENT APPLICATION NUMBER: US/09/347, 001A  
; CURRENT FILING DATE: 1999-07-02  
; EARLIER APPLICATION NUMBER: JP 10-193472  
; EARLIER FILING DATE: 1998-07-08  
; EARLIER APPLICATION NUMBER: JP 10-310398  
; EARLIER FILING DATE: 1998-10-30  
; EARLIER APPLICATION NUMBER: JP 11-12244  
; EARLIER FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1438  
; TYPE: DNA  
; ORGANISM: Unknown Organism

```

FEATURE:
;
; OTHER INFORMATION: Description of Unknown Organism: strain P528
;
; FEATURE:
;
; OTHER INFORMATION: N at position 1365 is A, T, G, or C
US-09-347-001-1

Query Match      58.3%; Score 849; DB 3; Length 1438;
Best Local Similarity 77.8%; Pred. No. 1,8e-282;
Matches 1136; Conservative 0; Mismatches 266; Indels 58; Gaps 7

Db      4  GAAACCTGGCGGCATGCTTTTACATGCAATGCAAGTCGAACCGGCAGACGAGATGCTTGATCTG 63
      18  GAACCTGGCGGCATGCTTTACACATGCAAGTCGACGSA-----CCTTTC 63

Qy      64  GTGGCAGATGGCGGACGGGTGAGTAATGATCGAATCGTATCGAAGAGGGGGTATACG 123
Db      64  GGGGTGATGGGCGGACGGGTGAGTAACGGGTAGGATCTATCCAAGGGGGGATTAACA 123

Qy      124  CATGAAAGATGTGCTAATACCGCATATCTTAAGAGAAAGACAGGGGATCGAAAGAC 183
Db      124  CTGGGAAACTGGTGTCTAATACCGCATATGACTGTAGGGGTCAAGGCGCGAGT----- 175

Qy      184  CTWGGCTTTTGGAGCGGCGCATGTCGTGATTAGTACTAGTTGGTGGGGTAAAGGCTTACAA 243
Db      176  ----CGCCTGTGAGAGAGCTTCGCTTCATTTAGCTTTGGTGGGGTAAAGGCTTACAA 231

Qy      244  GCGCAGATCAAGTATGGTGTGAGAGGACGACCAAGCCACACTGGGAGCTGAGACAGGCTC 303
Db      232  GCGCATGATCGATGACTGTGTGAGAGGATGATCGACCACACTGGGACTGAGACAGGCTC 291

Qy      304  CAGATCTCTTACGGGAGGCGACAGTGGGGAATTTTGGACATGGGGCGCAGCTGATCCAG 363
Db      292  CAGATCTCTTACGGGAGGCGACAGTGGGGAATTTTGGACATGGGGCGCAGCTGATCCAG 351

Qy      364  CAATGCGCGTGAAGGAAAGAGGAGCCCTTCGGGTTGTAAAGCTCTTTCAGTGCAGAAAGAAA 423
Db      332  CAATGCGCGTGTGTGAAGAGAGGCTTTGGGATTTGAAAGCACTTTGACAGGGAGCA---- 408

Qy      424  GGTTCAGGTAAATAATTCGTGACTCATGACGGTATCGACAGAAAGAACCGGCTTACTATC 483
Db      409  -----TGATGACGGTATCCCGTATGAGAAAGCCCGGCTTACTATC 446

Qy      484  GTGCCAGACAGCCGCGGTAAATACGTATGGGTGCAGAGGCTTAATGTGGAATTACTGGGCTTAA 543
Db      447  GTGCCAGACAGCCGCGGTAAATACGAAGGGGGGCTAGCGGTGCTCGAATGACTGGGCTTAA 506

Qy      544  GGGTGCAGAGGCGGCTTTGTAAAGTACAGATGTGAATATCCCGGGCTTAACTGGGAATTGC 603
Db      507  GGGCTGTGAGGCGGCTTTGTAAAGTACAGATGTGAATATCCAGGGCTTAACTGGGGCTGC 566

Qy      604  GTTTGAAACTCAAGGCTTAAAGTGTGGCAGAGAGGAGTGGAAATTCATGTGTAGCAGTGA 663
Db      567  ATTTGATATCGTAGCGACTTAGATGTGAGAGAGGTTGTGGAAATTCAGATGTATAGAGTGA 626

Qy      664  AATGCGTATGAGATATGAAAGAACATCGATGCGGAAAGGCGAGCTCCTGGGTATCAACTGAC 723
Db      627  AATTCGTATGATATTTGGGAAAGAACACCGGTGGGAAAGGCGGCAACTGGTCTCATGACTGAC 686

Qy      724  GCTCATGACAGAAACGTGGGGAGCAACAGGATTAAGTATCCCTGTGTGTCCAGCGCCCTTA 783
Db      687  GTTGAGGCGCGAAACGTGGGGAGCAACAGGATTAAGTATCCCTGTGTGTCCAGCGCTTGA 746

Qy      784  AACGATGTCAACTAAGTTGTGGGCTTATTAAG--CTTGGTAAAGAACTTAACGCGCTGAAG 842
Db      747  AACGATGTGTCTGATATGTGGGTAACTTAAGTATCTCAGTGTGAAAGCTTAAGCGCGCTAAG 806

Qy      843  TTGACCGCTTGGGAGTACCGTTCGCAAGATTTAAATCTCAAGGAATTGACGGGACCCGC 902
Db      807  CACACCGCTTGGGAGTACCGGCTCGCAAGTTGAAACTCAAGGAATTGACCGGAGCCCGC 866

Qy      903  ACAAGCGGTGATTAATGTGAGTAAATTTGATGCCAAGGAAATACCTTAACCTAACCTTGA 962
Db      867  ACAACCGGTGAGCAATGTGTAAATTTGAAAGCAACGCGCAAGACCTTACAGGGCTTGA 926

```



Db 964 UUACGAGGUCUUGACAUCCUUGACCAUCUUGAGACAGACUCCUUGCGGAGGCAAA 1023  
 Qy 1008 AACACAGGTCGTGACGTCGTGACGTCGTGTCGTGATGTTGGTTAAGTCCGC 1067  
 Db 1024 GUGACAGGUGGUGCAUGGUGCGUCACUCCUGUGUAGUUGGUGUUAAGCCCGC 1083  
 Qy 1068 AACGAGCGCAACCCCTTGATTAATGTCATCATTTGGTGGCACTTTAAAGACATGC 1127  
 Db 1084 AACGAGCGCAACCCUUGAUCUUAAGUCCAGCAUUAAGUUGGCAUCUUAAGUACUC 1143  
 Qy 1128 CGGTGACAAACCGAGAGAAAGTGGGATGACGTAAGTCTCAATGAGCCCTTAAGGATAG 1187  
 Db 1144 CGUGACAAACCGAGAGAAAGTGGGATGACGTAAGTCTCAATGAGCCCTTAAGGATAG 1203  
 Qy 1188 GTTACACAGTAAATCAATGCGCGCTACAGAGGTTTCCAAACCGGAGGGGAGACTAT 1247  
 Db 1204 GCUACACAGUCUUAAGUAGUAGUACAAAGGUGCGAAGCCGAGAGUUAACCAU 1263  
 Qy 1248 CTGAGAAAGCGGTCGTAGTCCGATCGAGTCTGCACTCGAATCCGTAAGTGGAAAT 1307  
 Db 1264 CCCAUAAGCCAUUCUACAUUGGAGUUGCAGGCGUACUCCCGUAGAAAGCCGAU 1323  
 Qy 1308 CGCTAGTAATCGCGATGACATGTCGCGTGAATACTTCCGAGCTTGTGACACCG 1367  
 Db 1324 UGCUAGUAUCCGCGAUACAGACGCGCGUGAUAUCCGCGGUCUUGUACACACG 1383  
 Qy 1368 CCCGTCACACCATGGAGTGGTTTTCACAGAAAGCAGTACTTAACCTAAGAGAGGCG 1427  
 Db 1384 CCCGTCACACCAAGAGUUGUUAACCCGAAAGCGUGAGUUAACCUUUGAGAGCG 1443  
 Qy 1428 CTGGCAGCGTGAATTCATGATGAGGAGT 1457  
 Db 1444 CCGCUAAGUGGAGCAUAUUGGUGUG 1473

RESULT 68  
 US-09-694-531-13  
 : Sequence 13, Application US/09694531  
 : Patent No. 6368843  
 : GENERAL INFORMATION:  
 : APPLICANT: Andersen, Lene N.  
 : APPLICANT: Schulten, Martin  
 : APPLICANT: Lange, Niels E.  
 : APPLICANT: Bjornvad, Made E.  
 : APPLICANT: Moller, Soren  
 : APPLICANT: Glad, Sanne O. S.  
 : APPLICANT: Kauppinen, Markus S.  
 : APPLICANT: Schnotz, Kirk  
 : APPLICANT: Kongsbak, Lars  
 : TITLE OF INVENTION: No. 6368843el Pectate Lyases  
 : FILE REFERENCE: 5378.200-US  
 : CURRENT APPLICATION NUMBER: US/09/694,531  
 : PRIOR FILING DATE: 2000-10-23  
 : PRIOR APPLICATION NUMBER: 09/198,955  
 : PRIOR FILING DATE: 1998-11-24  
 : PRIOR APPLICATION NUMBER: 1343/97  
 : PRIOR FILING DATE: 1997-11-24  
 : PRIOR APPLICATION NUMBER: 1344/97  
 : PRIOR FILING DATE: 1997-11-24  
 : PRIOR APPLICATION NUMBER: 60/067,249  
 : PRIOR FILING DATE: 1997-12-02  
 : PRIOR APPLICATION NUMBER: 09/073,684  
 : PRIOR FILING DATE: 1998-05-06  
 : PRIOR APPLICATION NUMBER: 09/184,217  
 : PRIOR FILING DATE: 1998-11-02  
 : NUMBER OF SEQ ID NOS: 32  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 13  
 : LENGTH: 1506  
 : TYPE: RNA  
 : ORGANISM: Bacillus sp. 1534

US-09-694-531-13  
 Query Match 58.2%; Score 848.4; DB 3; Length 1506;  
 Best Local Similarity 61.7%; Pred. No. 3e-282;  
 Matches 907; Conservative 216; Mismatches 331; Indels 16; Gaps 6;  
 Qy 4 GAAAGTGGCGGCAATGCTTTACATGACAGTCCAGCGGACGATGCTTCATCTG 63  
 Db 4 GAAAGTGGCGGCGGCGGCUUAUAUACUAGCAAGUCCAGGAGUUAAGGAGGAGUUCUCCG 63  
 Qy 64 GTGCGAGTGGCGGACCGGAGTGAATGATCGG-AACTGATCCAGAAAGGCGGTTAAC 122  
 Db 64 GAUUAUUGCGCGGACCGGAGUUAACAAGGCGCAACUUGCCUUAAGACUUGGAGUAA 123  
 Qy 123 GCATCGAAAGATGTGCTAATACCGCAT-ATATCTTAAGAGAAAGCAAGGAGATCGAA 181  
 Db 124 UCCGGGAAACCGGUCUUAUACCGAAUAAACUUGAACCUUGGUGUCAAAGUUAAG 183  
 Qy 182 A-----CTTGGCTTTTGAAGCGGCGGATGTCTGATTAATGTTGTTGTTGTTGTTGTT 231  
 Db 184 AUGGCUUGGUCUUAUACUUAAGAAUGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 243  
 Qy 232 AAGGCTTACCAAGGCGAGATCAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 291  
 Db 244 AUGGCUUAACCAAGGCGAGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 303  
 Qy 292 TGAACACGCGGCGGACCTCTTACCGGAGGACGAGTGGGAAATTTTGAACAAATGGGCGCA 351  
 Db 304 UGAACACGCGGCGGACCTCTTACCGGAGGACGAGTGGGAAATTTTGAACAAATGGGCGCA 363  
 Qy 352 AGCTGATCCAGCAATGCGCGGATGAGTGAAGAAAGGCTTGGGTTGTTGTTGTTGTTGTTGTT 411  
 Db 364 AGUUGACGAGCAACCGCGGUGAGUAGAGAAAGCCUUGCGGUGUAAAGUUCUUGU 423  
 Qy 412 TCGAGAAAGAAAGGTTACGG-TAATATCTGATCTGATGACGTTATCGACAGAAAGC 470  
 Db 424 GAGGGAAGAAACAAAGUUAUCCGUGAUAUAAAGCGUACCUUAGACGUAACCAAGAAAGC 483  
 Qy 471 ACCGCTTAATCTGCTGACGACGCGGCTTAATCTGATGTTGTTGTTGTTGTTGTTGTTGTT 530  
 Db 484 CACGCUAAACUACGUGCACGACGCGGUAUAUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 543  
 Qy 531 TACTGGCGTAAAGGAGGCGGAGGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 590  
 Db 544 UAUUGGCGUAAAGCGCGGAGGCGGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 603  
 Qy 591 ACTGGGAATTCCTTTGAAACTCAAGGCTAGTGTGACAGAGGAGTGAATTTCA 650  
 Db 604 ACCCGAGCGGCGUAGAAACUUGGAGGCUUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 663  
 Qy 651 TGTGTACGATGAATTCGTTAGATATGGAAGAACTGATGCGGACGACGCTCTCG 710  
 Db 664 CGUUAAGCGUAGAAUUGCGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 723  
 Qy 711 GGTAACTGACGCTGATGACGAAAGCTGGGAGCAACAGATGATATGATACCTCTGT 770  
 Db 724 GUCUGUAACUAGCGCUAGAGCGGAAAGCGUGGAGCAACAGAUUAUAGUAGUAGUAGUAGU 783  
 Qy 771 AGTCAAGCCCTTAACGATGCTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 828  
 Db 784 AGUCCAGCGGUAACGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 843  
 Qy 829 GCTAACGCGTGAAGTTAAGCGGCTGGGAGTACGTTGCGAAATTAATACTCAAGGAAT 888  
 Db 844 GUAAACAUUAAGCAUCGCGCGGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 903  
 Qy 889 TGAAGGGAACCGGACAAAGCGGTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 948  
 Db 904 UAGGAGGAGACCGGACAAAGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 963  
 Qy 949 TTACTTACCTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1007  
 Db 964 UUAACGAGUUCUUAACAUCCUUAACCAUCUUGAGACAGAGCUUCCUUGCGGAGGCAAA 1023









**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 5, 2006, 13:14:32 ; Search time 1262 Seconds  
(without alignments)  
9547.135 Million cell updates/sec

Title: US-10-659-980A-1

Perfect score: 1457  
Sequence: 1 atgacacgtgctggcgatgc.....tgaatcatgactggggtc 1457

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 70 summaries

Database : Published Applications NA Main:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1457	100.0	1457	US-10-659-948A-1	Sequence 1, Appl
2	1457	100.0	1457	US-10-659-980A-1	Sequence 1, Appl
3	1457	100.0	1457	US-10-659-983A-1	Sequence 1, Appl
4	1452.2	99.7	1457	US-10-659-948A-2	Sequence 2, Appl
5	1452.2	99.7	1457	US-10-659-980A-2	Sequence 2, Appl
6	1452.2	99.7	1457	US-10-659-983A-2	Sequence 2, Appl
7	1391.6	95.5	1457	US-10-659-948A-20	Sequence 20, Appl
8	1391.6	95.5	1457	US-10-659-980A-20	Sequence 20, Appl
9	1391.6	95.5	1457	US-10-659-983A-20	Sequence 20, Appl
10	1359.2	93.3	1457	US-10-659-948A-19	Sequence 19, Appl
11	1359.2	93.3	1457	US-10-659-980A-19	Sequence 19, Appl
12	1359.2	93.3	1457	US-10-659-983A-19	Sequence 19, Appl
13	1358.8	93.3	1457	US-10-659-948A-18	Sequence 18, Appl
14	1358.8	93.3	1457	US-10-659-980A-18	Sequence 18, Appl
15	1358.8	93.3	1457	US-10-659-983A-18	Sequence 18, Appl
16	1287.6	88.4	1458	US-10-659-948A-3	Sequence 3, Appl
17	1287.6	88.4	1458	US-10-659-980A-3	Sequence 3, Appl
18	1287.6	88.4	1458	US-10-659-983A-3	Sequence 3, Appl
19	1230.2	84.4	1460	US-10-659-980A-4	Sequence 4, Appl
20	1230.2	84.4	1460	US-10-659-983A-4	Sequence 4, Appl
21	1230.2	84.4	1460	US-10-659-980A-4	Sequence 4, Appl
22	1206.4	82.8	1485	US-10-168-337A-10	Sequence 10, Appl
23	1194	81.9	1460	US-10-168-337A-6	Sequence 6, Appl

## ALIGNMENTS

24	1181	81.1	1459	6	US-10-168-337A-8	Sequence 8, Appl
25	1179.2	80.9	1460	6	US-10-168-337A-4	Sequence 4, Appl
26	1156.8	79.4	1460	6	US-10-168-337A-2	Sequence 2, Appl
27	1155.8	79.3	1460	6	US-10-168-337A-1	Sequence 1, Appl
28	1152.4	79.1	1532	6	US-10-029-357A-38	Sequence 38, Appl
29	1143.8	78.5	1496	6	US-10-723-365B-32	Sequence 32, Appl
30	1142	78.4	1426	6	US-10-168-337A-9	Sequence 9, Appl
31	1138.4	78.1	1535	6	US-10-029-337A-41	Sequence 41, Appl
32	1137.6	78.1	1530	8	US-10-875-161-1	Sequence 1, Appl
33	1137.2	78.1	1485	6	US-10-029-337A-39	Sequence 39, Appl
34	1131.4	77.7	1464	6	US-10-029-337A-40	Sequence 40, Appl
35	1130.6	77.6	1478	6	US-10-168-337A-3	Sequence 3, Appl
36	1129.2	77.5	1481	6	US-10-397-551-1	Sequence 1, Appl
37	1129.2	77.5	1481	10	US-11-112-257-1	Sequence 1, Appl
38	1126.8	77.2	1488	6	US-10-029-337A-42	Sequence 42, Appl
39	1124.6	77.2	1544	6	US-10-029-337A-45	Sequence 45, Appl
40	1124.6	77.2	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap
41	1124.6	77.2	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap
42	1122	77.0	1610	6	US-10-029-337A-43	Sequence 43, Appl
43	1119.8	76.9	1544	3	US-09-726-774-5	Sequence 5, Appl
44	1119.8	76.9	1544	6	US-10-029-337A-44	Sequence 44, Appl
45	1119.8	76.9	1544	7	US-10-719-633-5	Sequence 5, Appl
46	1108	76.0	1451	6	US-10-219-549-1	Sequence 1, Appl
47	1103.2	75.7	1453	6	US-10-464-724-4	Sequence 4, Appl
48	1103.2	75.7	1453	6	US-10-464-356-4	Sequence 4, Appl
49	1103.2	75.7	1453	6	US-10-464-709-4	Sequence 4, Appl
50	1103	75.7	1440	6	US-10-678-023-1	Sequence 1, Appl
51	1100.6	75.5	20844	9	US-10-915-740A-8	Sequence 8, Appl
52	1089.4	74.8	1322	6	US-10-168-337A-5	Sequence 5, Appl
53	1080.8	74.2	1539	6	US-10-464-356-6	Sequence 6, Appl
54	1080.8	74.2	1539	6	US-10-464-709-5	Sequence 5, Appl
55	1080.8	74.1	1534	6	US-10-464-724-7	Sequence 7, Appl
56	1080.2	74.1	1454	6	US-10-464-356-7	Sequence 7, Appl
57	1080.2	74.1	1454	6	US-10-464-709-7	Sequence 7, Appl
58	1080.2	74.1	1454	6	US-10-464-724-2	Sequence 2, Appl
59	1079.6	74.1	1467	6	US-10-464-356-2	Sequence 2, Appl
60	1079.6	74.1	1467	6	US-10-464-709-2	Sequence 2, Appl
61	1079.6	74.1	1467	6	US-10-168-337A-7	Sequence 7, Appl
62	1079.2	74.1	1320	6	US-09-985-846-1	Sequence 1, Appl
63	1069.8	72.8	1495	3	US-10-776-761-1	Sequence 1, Appl
64	1060.8	72.3	1540	3	US-09-967-376-1	Sequence 1, Appl
65	1053	72.3	1540	5	US-10-260-647-1	Sequence 1, Appl
66	1053	72.3	1540	5	US-10-260-647-1	Sequence 1, Appl
67	1041.8	71.5	2731748	7	US-10-297-465A-1	Sequence 53, Appl
68	1031.8	70.4	1547	6	US-10-418-861B-53	Sequence 3, Appl
69	1026.4	70.4	1450	6	US-10-464-724-3	Sequence 3, Appl
70	1026.4	70.4	1450	6	US-10-464-356-3	Sequence 3, Appl

RESULT 1  
US-10-659-948A-1  
Sequence 1, Application US/10659948A  
Publication No. US20040101946A1  
GENERAL INFORMATION:  
APPLICANT: Hovancg, Timothy A  
TITLE OF INVENTION: Method of using Ammonia-Oxidizing Bacteria  
FILE REFERENCE: 81289-294309  
CURRENT APPLICATION NUMBER: US/10/659, 948A  
PRIOR FILING DATE: 2003-09-10  
PRIOR APPLICATION NUMBER: US 09/573, 684  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 60/386, 217  
PRIOR FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: US 60/386, 218  
PRIOR FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: US 60/386, 219  
PRIOR FILING DATE: 2002-09-19  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1

LENGTH: 1457  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: AOB Type A R7c1one140 16S rDNA  
US-10-659-948A-1

Query Match 100.0%; Score 1457; DB 7; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 1 ATTGAACGCTGCGGCATGCTTTTACATGCAAGTCGAACGCGACGAGATGCTTGAT 60
Db 1 ATTGAACGCTGCGGCATGCTTTTACATGCAAGTCGAACGCGACGAGATGCTTGAT 60
Oy 61 CTGTGGCGAGTGGCGGACGCGGTGATATGATCGAAGCGTATCCAGAAGAGGGGGTA 120
Db 61 CTGTGGCGAGTGGCGGACGCGGTGATATGATCGAAGCGTATCCAGAAGAGGGGGTA 120
Oy 121 ACGCATCGAAGATGCTTAATACCGCATATCTCTAAGAGGAAGAGGGGATCGAAA 180
Db 121 ACGCATCGAAGATGCTTAATACCGCATATCTCTAAGAGGAAGAGGGGATCGAAA 180
Oy 181 GACCTTGGCTTTTGGACGCGCGCATGCTGATTAAGTGTGGGTAAAGGCTTAC 240
Db 181 GACCTTGGCTTTTGGACGCGCGCATGCTGATTAAGTGTGGGTAAAGGCTTAC 240
Oy 241 CAAGCGCATGATCAAGTGTGCTGAGAGACGACCGCATCTGGGACTGAGACACG 300
Db 241 CAAGCGCATGATCAAGTGTGCTGAGAGACGACCGCATCTGGGACTGAGACACG 300
Oy 301 GCCCAGATCTCTACCGGAGGACGAGTGGGGAAATTTGGAACAATGGGGCAAGCTGATC 360
Db 301 GCCCAGATCTCTACCGGAGGACGAGTGGGGAAATTTGGAACAATGGGGCAAGCTGATC 360
Oy 361 CAGCAATCCCGCTGAGTGAAGAAAGCCCTCGGGTGTAAAGCTTTTCAATCGAGAGA 420
Db 361 CAGCAATCCCGCTGAGTGAAGAAAGCCCTCGGGTGTAAAGCTTTTCAATCGAGAGA 420
Oy 421 AAAGGTTACGGTAAATTAATCGTACTCATGACGGTATGACAGAAAGACCGGCTAAC 480
Db 421 AAAGGTTACGGTAAATTAATCGTACTCATGACGGTATGACAGAAAGACCGGCTAAC 480
Oy 481 TACGTCGACGACGCGCGTAAATGATGAGGGTGAAGGTTATGGAATTAATCTGGGGT 540
Db 481 TACGTCGACGACGCGCGTAAATGATGAGGGTGAAGGTTATGGAATTAATCTGGGGT 540
Oy 541 AAAGGTCGACGAGCGGCTTTGTAAGTCAAGATGTGAATCCCGGGCTTAACTTGGGAT 600
Db 541 AAAGGTCGACGAGCGGCTTTGTAAGTCAAGATGTGAATCCCGGGCTTAACTTGGGAT 600
Oy 601 TGCCTTTGAAAATCAAGAGCTAGAGTGTGCGCAGAGGGAGTGAATTCATGTTAGAGAG 660
Db 601 TGCCTTTGAAAATCAAGAGCTAGAGTGTGCGCAGAGGGAGTGAATTCATGTTAGAGAG 660
Oy 661 TGAATGGTGAAGTATGGAAGAACATGATGGGAGAGGCGCTCTGGGTTAACT 720
Db 661 TGAATGGTGAAGTATGGAAGAACATGATGGGAGAGGCGCTCTGGGTTAACT 720
Oy 721 GACGCTCATGACGAAACGCTGGGAGCAACAGATTAGATCCCTGATGTCACGCGC 780
Db 721 GACGCTCATGACGAAACGCTGGGAGCAACAGATTAGATCCCTGATGTCACGCGC 780
Oy 781 CTAAACGATGTCACATGTTGTTGGGCTTATTAGGCTTGTGAACGAAGCTAACCGGTGA 840
Db 781 CTAAACGATGTCACATGTTGTTGGGCTTATTAGGCTTGTGAACGAAGCTAACCGGTGA 840
Oy 841 AGTTGACCGGCTGGGAGATGAGTGGCAAGATTAACCAAGGAATTGACGGGGAACC 900
Db 841 AGTTGACCGGCTGGGAGATGAGTGGCAAGATTAACCAAGGAATTGACGGGGAACC 900
Oy 901 GCACAAAGCGGTGATTAATGTGGAATTAATTCATGCAACGCGAATAACCTTAACCTT 960
Db 901 GCACAAAGCGGTGATTAATGTGGAATTAATTCATGCAACGCGAATAACCTTAACCTT 960
```

```
Db 901 GCACAAAGCGGTGATTAATGTGGAATTAATTCATGCAACGCGAATAACCTTAACCTT 960
Oy 961 GACATGTAAGCAATTTTCTAGAGTATAGTCTTCGGGAAAGCTTAACAGAGTGTG 1020
Db 961 GACATGTAAGCAATTTTCTAGAGTATAGTCTTCGGGAAAGCTTAACAGAGTGTG 1020
Oy 1021 CATGGCTGTGCTCAAGCTCGTGTGAGATGTTAGTGGGTTAAGTCCCGCAAGAGGCAACC 1080
Db 1021 CATGGCTGTGCTCAAGCTCGTGTGAGATGTTAGTGGGTTAAGTCCCGCAAGAGGCAACC 1080
Oy 1081 CTGTGATTAATTCATCATATTTGTTGGGCACTTTAATGAGATCGCCGCTGACAAACCG 1140
Db 1081 CTGTGATTAATTCATCATATTTGTTGGGCACTTTAATGAGATCGCCGCTGACAAACCG 1140
Oy 1141 GAGGAAGTGGGAGTATGATGCTCAAGTCCATGAGCCCTTAATGGGATTCACACGTTA 1200
Db 1141 GAGGAAGTGGGAGTATGATGCTCAAGTCCATGAGCCCTTAATGGGATTCACACGTTA 1200
Oy 1201 TACAATGGCGGTACAGAGGTTGCCAACCCGCAAGGGGAGCTTAATCTCAAAAAGCGCG 1260
Db 1201 TACAATGGCGGTACAGAGGTTGCCAACCCGCAAGGGGAGCTTAATCTCAAAAAGCGCG 1260
Oy 1261 TCGTAGTCCGATCGAGTCTGCAACTCGACTCCGTGAAGTCGGAATCGCTAGTATGCG 1320
Db 1261 TCGTAGTCCGATCGAGTCTGCAACTCGACTCCGTGAAGTCGGAATCGCTAGTATGCG 1320
Oy 1321 GGATCAGATGTCGCGGTAATAGTTCGCGGGTCTTGATACACCGCGCGCTCACACCAT 1380
Db 1321 GGATCAGATGTCGCGGTAATAGTTCGCGGGTCTTGATACACCGCGCGCTCACACCAT 1380
Oy 1381 GGGAGTGGGTTTCAACGAAGCAGATGCTTAACCGTAAGAGAGGCGCTTGCCACGCTGA 1440
Db 1381 GGGAGTGGGTTTCAACGAAGCAGATGCTTAACCGTAAGAGAGGCGCTTGCCACGCTGA 1440
Oy 1441 GATTCAATGATCGGGGTG 1457
Db 1441 GATTCAATGATCGGGGTG 1457
```

RESULT 2  
US-10-659-980A-1  
; Sequence 1, Application US/10659980A  
; Publication No. US20040106133A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovaneq, Timothy A  
; TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-284781  
; CURRENT APPLICATION NUMBER: US/10/659, 980A  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573, 684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386, 217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386, 218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386, 219  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 1457  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: AOB Type A R7c1one140 16S rDNA  
US-10-659-980A-1

Query Match 100.0%; Score 1457; DB 7; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 1 ATTGAACGCTGCGGCATGCTTTTACATGCAAGTCGAACGCGACGAGATGCTTGAT 60
Db 1 ATTGAACGCTGCGGCATGCTTTTACATGCAAGTCGAACGCGACGAGATGCTTGAT 60
```

```
Db 1 ATTGAACGCTG3CGGCATGCTTTACACATGCAAGTCGAACGCGACGACGATGCTTGAT 60
Qy 61 CTGGGCGAGTGGCGGACCGGGTGAATGATCATCGAACGTAATCCAGAMAGGGGGGTA 120
Db 61 CTGGGCGAGTGGCGGACCGGGTGAATGATCATCGAACGTAATCCAGAMAGGGGGGTA 120
Qy 121 ACGCATCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAA 180
Db 121 ACGCATCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAA 180
Qy 181 GACCTTGCGCTTTTGGAGCGGCCATGTCTGATTAGCTAAGTTGGTGGGTAAGGCGTAC 240
Db 181 GACCTTGCGCTTTTGGAGCGGCCATGTCTGATTAGCTAAGTTGGTGGGTAAGGCGTAC 240
Qy 241 CAAGCGCAGATCACTAGTGGTCTGAGAGACGACCACTGGGACTGAGACAG 300
Db 241 CAAGCGCAGATCACTAGTGGTCTGAGAGACGACCACTGGGACTGAGACAG 300
Qy 301 GCCCAGACTCCTTACGGGAGGACGAGTGGGAAATTTGGACATAGGCGCAAGCTTGATC 360
Db 301 GCCCAGACTCCTTACGGGAGGACGAGTGGGAAATTTGGACATAGGCGCAAGCTTGATC 360
Qy 361 CAGCATATCCCGCTGAGTGAAGAGAGCCCTTGGGGTTGTAAGCTCTTTCAGTCGAGAA 420
Db 361 CAGCATATCCCGCTGAGTGAAGAGAGCCCTTGGGGTTGTAAGCTCTTTCAGTCGAGAA 420
Qy 421 AAAAGTTACGGTAATTAATCGTACATGACGATGACGAGTAAGACAGAAAGACCGGCTAAC 480
Db 421 AAAAGTTACGGTAATTAATCGTACATGACGATGACGAGTAAGACAGAAAGACCGGCTAAC 480
Qy 481 TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGCTTAATCGGAATTACTGCGCGT 540
Db 481 TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGCTTAATCGGAATTACTGCGCGT 540
Qy 541 AAAAGGTGGCGACGGCGCTTTGTAAGTCAAGTGTAAATCCCGGGCTTAACCTGGGAT 600
Db 541 AAAAGGTGGCGACGGCGCTTTGTAAGTCAAGTGTAAATCCCGGGCTTAACCTGGGAT 600
Qy 601 TGGCGTTGAAACTAACAAGGCTAGAGTGTGGAGAGGAGGAGTGAATTCATGTGTAGCAG 660
Db 601 TGGCGTTGAAACTAACAAGGCTAGAGTGTGGAGAGGAGGAGTGAATTCATGTGTAGCAG 660
Qy 661 TGAATGTGCTAGATATGAGAGAAATCATGATGCGAGCAAGCCCTCTGGGTTAACT 720
Db 661 TGAATGTGCTAGATATGAGAGAAATCATGATGCGAGCAAGCCCTCTGGGTTAACT 720
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGGTATGCCAGCC 780
Db 721 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGGTATGCCAGCC 780
Qy 781 CTAAACGATGTCAACTAGTGTGGGCTTAATTAAGGCTTGTAAAGAGCTAAACGCGTGA 840
Db 781 CTAAACGATGTCAACTAGTGTGGGCTTAATTAAGGCTTGTAAAGAGCTAAACGCGTGA 840
Qy 841 AGTTGACCGCTGGGAGATACGCTCGCAAGATTTAAACTCAAGAAATGAACGGGAGCC 900
Db 841 AGTTGACCGCTGGGAGATACGCTCGCAAGATTTAAACTCAAGAAATGAACGGGAGCC 900
Qy 901 GCAACAAGCGGTGATATGTGATTAATTCGATGCAACGCGGAAAACTTTACCTACCTT 960
Db 901 GCAACAAGCGGTGATATGTGATTAATTCGATGCAACGCGGAAAACTTTACCTACCTT 960
Qy 961 GACATGTACGGAATTTTCTAGAGATAGATTAGTTCGGGAAAGGCTAACACAGGTGCTG 1020
Db 961 GACATGTACGGAATTTTCTAGAGATAGATTAGTTCGGGAAAGGCTAACACAGGTGCTG 1020
Qy 1021 CATGCGTGTGTCAGCTCGTGTGTGAGATGTTGGTTAAGTCCCGCAAGCAGCGCAACC 1080
Db 1021 CATGCGTGTGTCAGCTCGTGTGTGAGATGTTGGTTAAGTCCCGCAAGCAGCGCAACC 1080
Qy 1081 CTGTGATTAATTTGCAATCATTTGTTGGCACTTTAATGAGACTGCGGCTGACAAACCG 1140
Db 1081 CTGTGATTAATTTGCAATCATTTGTTGGCACTTTAATGAGACTGCGGCTGACAAACCG 1140
```

```
Qy 1141 GAGGAGGTGGGATGAGCTCAAGTCCCTAATGCGCTTAATGGGTAGGAGCTTCAACGTTAA 1200
Db 1141 GAGGAGGTGGGATGAGCTCAAGTCCCTAATGCGCTTAATGGGTAGGAGCTTCAACGTTAA 1200
Qy 1201 TACAATGGCGGTCAAGAGGTTGCCAACCCGCGAGGGGAGTAACTCAGAAAGCGCG 1260
Db 1201 TACAATGGCGGTCAAGAGGTTGCCAACCCGCGAGGGGAGTAACTCAGAAAGCGCG 1260
Qy 1261 TCGTAGTCCGGATCGAGTCTGCAACTCGCTGTAAGTCCGGAATGCTAAGTAAATGCG 1320
Db 1261 TCGTAGTCCGGATCGAGTCTGCAACTCGCTGTAAGTCCGGAATGCTAAGTAAATGCG 1320
Qy 1321 GGATCAGATATGCTCCGCGTAATACGTTCCCGGGCTTGTACACCGCCCGTCAACCAT 1380
Db 1321 GGATCAGATATGCTCCGCGTAATACGTTCCCGGGCTTGTACACCGCCCGTCAACCAT 1380
Qy 1381 GGGAGTGGGTTTCAACAGAGAGTGTCTAACCGTAAAGAGAGGCGCTTGCACAGTGA 1440
Db 1381 GGGAGTGGGTTTCAACAGAGAGTGTCTAACCGTAAAGAGAGGCGCTTGCACAGTGA 1440
Qy 1441 GATTCAATGACTGGGCTG 1457
Db 1441 GATTCAATGACTGGGCTG 1457

RESULT 3
US-10-659-983A-1
; Sequence 1, Application US/10659983A
; Publication No. US2004015713A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneq, Timothy A
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-28479
; CURRENT APPLICATION NUMBER: US/10/659,983A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type A R7clome140 16S rDNA
US-10-659-983A-1

Query Match 100.0%; Score 1457; DB 7; Length 1457;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATTGAACGCTGCGGAGATCTTAACATGCAAGTCGAACGCGACGACGATGCTTGAT 60
Db 1 ATTGAACGCTGCGGAGATCTTAACATGCAAGTCGAACGCGACGACGATGCTTGAT 60
Qy 61 CTGGTGGCAGTGGCGGACCGGGTGAATGATCATCGAACGTAATCCAGAMAGGGGGGTA 120
Db 61 CTGGTGGCAGTGGCGGACCGGGTGAATGATCATCGAACGTAATCCAGAMAGGGGGGTA 120
Qy 121 ACGCATCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAA 180
Db 121 ACGCATCGAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAA 180
Qy 181 GACCTTGCGCTTTTGGAGCGGCCATGTCTGATTAGCTAAGTTGGTGGGTAAGGCGTAC 240
Db 181 GACCTTGCGCTTTTGGAGCGGCCATGTCTGATTAGCTAAGTTGGTGGGTAAGGCGTAC 240
```

241 CAAGGCGAGCATCAGTAAGTGGTCTGAGAGGAGCAAGCCGCACTGGGAGCTGAGACACG 300  
241 CAAGGCGAGCATCAGTAAGTGGTCTGAGAGGAGCAAGCCGCACTGGGAGCTGAGACACG 300  
301 GCCCAGACTCTTACCGGAGGAGCAGCAGTGGGGAATTTTGGACATAGGGGCGCAGCCTGTATC 360  
301 GCCCAGACTCTTACCGGAGGAGCAGCAGTGGGGAATTTTGGACATAGGGGCGCAGCCTGTATC 360  
361 CAGCAATGCCGCGTGAAGAGAGGCTTCCGGTTGTAAAGCTCTTTCAGTCGAGAAAGA 420  
361 CAGCAATGCCGCGTGAAGAGAGGCTTCCGGTTGTAAAGCTCTTTCAGTCGAGAAAGA 420  
421 AAAAGTTACGGTAAATTAATCCGTAATCCTCATGACGATGACGAGAAAGCAACCGGCTAAC 480  
421 AAAAGTTACGGTAAATTAATCCGTAATCCTCATGACGATGACGAGAAAGCAACCGGCTAAC 480  
481 TACGTCGACGAGCAGCCGCGGTAAATACGTAGGGTGCAGAGGCTTAATCGAATTAATCTGGCGT 540  
481 TACGTCGACGAGCAGCCGCGGTAAATACGTAGGGTGCAGAGGCTTAATCGAATTAATCTGGCGT 540  
541 AAAAGGTGCGCAGCGCGCTTTGTAAAGTCAAGATGTGAATCCCGGGCTTAACTGGGAAT 600  
541 AAAAGGTGCGCAGCGCGCTTTGTAAAGTCAAGATGTGAATCCCGGGCTTAACTGGGAAT 600  
601 TGGCGTTGAAACTCAAGGCTAGAGTGGGCGAGAGGGGTGGAAATTCATGTGTAGAG 660  
601 TGGCGTTGAAACTCAAGGCTAGAGTGGGCGAGAGGGGTGGAAATTCATGTGTAGAG 660  
661 TGAATGCGTGAAGATATGAGAGAAATCATGATGCGAAGGAGCCTCTGGGTTAACT 720  
661 TGAATGCGTGAAGATATGAGAGAAATCATGATGCGAAGGAGCCTCTGGGTTAACT 720  
721 GACGCTCATGCAAGAAAGCTGGGAGGAGCAACAGGATTAATGATACCTGTGATGCCAGCC 780  
721 GACGCTCATGCAAGAAAGCTGGGAGGAGCAACAGGATTAATGATACCTGTGATGCCAGCC 780  
781 CTTAAACGATGCAATAGTATGTTGGGCTTTAGGCTTGGTAAAGAAAGCTAACGCGGGA 840  
781 CTTAAACGATGCAATAGTATGTTGGGCTTTAGGCTTGGTAAAGAAAGCTAACGCGGGA 840  
841 AGTTGACCGCCTGGGAGGATACGCGTGCAGAGTTAAATCAAGAGAAATTTGAACGCGGAGCC 900  
841 AGTTGACCGCCTGGGAGGATACGCGTGCAGAGTTAAATCAAGAGAAATTTGAACGCGGAGCC 900  
901 GCAACAAGCGGTGAGATTATGTGGATTAAATTCGATGCAAGCGGAAAACCTTACCTTACCTT 960  
901 GCAACAAGCGGTGAGATTATGTGGATTAAATTCGATGCAAGCGGAAAACCTTACCTTACCTT 960  
961 GACATGTAGCGAAATTTTCTAGAGATAGATTAGTGGGGAAGCTTAAACAGGTCGCTG 1020  
961 GACATGTAGCGAAATTTTCTAGAGATAGATTAGTGGGGAAGCTTAAACAGGTCGCTG 1020  
1021 CATGGCTGTGCTCAGCTGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1080  
1021 CATGGCTGTGCTCAGCTGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1080  
1081 CTTGTGATTAATTCGATCAATTTGGTGGGCACTTTAATGAGACTGCGGTCGACAAACCG 1140  
1081 CTTGTGATTAATTCGATCAATTTGGTGGGCACTTTAATGAGACTGCGGTCGACAAACCG 1140  
1141 GAGGAGGTGGGAGATGAGTCAAGTCCATGATGCGGCTTAAATGGGTAGGGCTTACACGTTAA 1200  
1141 GAGGAGGTGGGAGATGAGTCAAGTCCATGATGCGGCTTAAATGGGTAGGGCTTACACGTTAA 1200  
1201 TACAATGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTTAATTCAGAAAACGCG 1260  
1201 TACAATGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTTAATTCAGAAAACGCG 1260  
1261 TCGTAGTCCGATCGGATCTTGCAACTGACTCCGTAAGTCGGAATGCTTAGTAATCGC 1320  
1261 TCGTAGTCCGATCGGATCTTGCAACTGACTCCGTAAGTCGGAATGCTTAGTAATCGC 1320

1321 GGATCAGCATGTCCGGGTGAATACGTTCCCGGGTCTTGTACACACCCCGCTCACACCAT 1380  
1321 GGATCAGCATGTCCCGGTGAATACGTTCCCGGGTCTTGTACACACCCCGCTCACACCAT 1380  
1381 GGGAGTGGGTTTACCAAGAACAGATAGCTTAACCGTAAAGAGGGGCTTGCACGCGTGA 1440  
1381 GGGAGTGGGTTTACCAAGAACAGATAGCTTAACCGTAAAGAGGGGCTTGCACGCGTGA 1440  
1441 GATTCAATGACTGGGGTG 1457  
1441 GATTCAATGACTGGGGTG 1457

RESULT 4  
US-10-659-948a-2  
; Sequence 2, Application US/10659948A  
; Publication No. US20040101946A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovaneq, Timothy A  
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-294309  
; CURRENT APPLICATION NUMBER: US/10/659, 948A  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573,684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386,217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,219  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 1457  
; TYPE: DNA  
; ORGANISM: Unknown  
; OTHER INFORMATION: MOB Type A1 R7clone187 16S rDNA  
US-10-659-948a-2

Query Match 99.7%; Score 1452.2; DB 7; Length 1457;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
1 ATTGAAGCTGGCGGCAATGCTTTACATATGCAAGTGCAGACGCGATGCTTGCAT 60  
1 ATTGAAGCTGGCGGCAATGCTTTACATATGCAAGTGCAGACGCGATGCTTGCAT 60  
61 CTGGTGGCAGTGGCGGAGCGGGTGAATGATCGGAAGCTATCCAGAAAGGGGGGTA 120  
61 CTGGTGGCAGTGGCGGAGCGGGTGAATGATCGGAAGCTATCCAGAAAGGGGGGTA 120  
121 ACGCATGAAAGATGTCTAATAACGCATTAATCTTAAGAGAGAAAGCAGGGATCGAAA 180  
121 ACGCATGAAAGATGTCTAATAACGCATTAATCTTAAGAGAGAAAGCAGGGATCGAAA 180  
181 GACCTGCGCTTTTGGAGCGCGGATGTCTGATTAAGCTAAGTGGTGGGTTAAAGGCTTAC 240  
181 GACCTGCGCTTTTGGAGCGCGGATGTCTGATTAAGCTAAGTGGTGGGTTAAAGGCTTAC 240  
241 CAAGGCGACATCAGTATGTTGCTGAGAGAGACAGACGACACACTGGGACTGAGACAG 300  
241 CAAGGCGACATCAGTATGTTGCTGAGAGAGACAGACGACACACTGGGACTGAGACAG 300  
301 GCCCAGACTCTTACCGGAGGAGCAGCAGTGGGGAATTTTGGACAAATGGGCGCAAGCCTGATC 360  
301 GCCCAGACTCTTACCGGAGGAGCAGCAGTGGGGAATTTTGGACAAATGGGCGCAAGCCTGATC 360  
361 CAGCAATGCCGCGTGAAGAGAGGCTTCCGGTTGTAAAGCTCTTTCAGTCGAGAAAGA 420  
361 CAGCAATGCCGCGTGAAGAGAGGCTTCCGGTTGTAAAGCTCTTTCAGTCGAGAAAGA 420





```
Db 601 TGCCTTAAACTCAAGAGCTAGAGTGTGCGAGAGGGAGGTGCAATTCATGTGTAGAG 660
Qy 661 TGAATGCGTAGAGATATGAAAGAAATCATGATGCGAAGGACCTCTCGGTTAACT 720
Db 661 TGAATGCGTAGAGATATGAAAGAAATCATGATGCGAAGGACCTCTCGGTTAACT 720
Qy 721 GACGCTCATGACCAAGCAACCGTGGGAGCAAGATTAAGATCCCTGTAGTCCAGGCC 780
Db 721 GACGCTCATGACCAAGCAACCGTGGGAGCAAGATTAAGATCCCTGTAGTCCAGGCC 780
Qy 781 CTTAAACGATGTCAACTAGTGTGGGCTTTATAGGCTTGGTAAAGCAAGCTAACGCGTGA 840
Db 781 CTTAAACGATGTCAACTAGTGTGGGCTTTATAGGCTTGGTAAAGCAAGCTAACGCGTGA 840
Qy 841 AGTTGACCCGCTGGGAGTAGCGGTGCAAGATTAAGATTAAGATTAAGATTAAGATTAAG 900
Db 841 AGTTGACCCGCTGGGAGTAGCGGTGCAAGATTAAGATTAAGATTAAGATTAAGATTAAG 900
Qy 901 GCACAAAGCGGTGATATATGATGATTAATTCGATGCAAGCGCAAAACCTTACCTTACCCTT 960
Db 901 GCACAAAGCGGTGATATATGATGATTAATTCGATGCAAGCGCAAAACCTTACCTTACCCTT 960
Qy 961 GACATGTAGCGAATTTTCTAGAGATAGATTAAGTGTGCGGAAACGCTAACACAGGTGCTG 1020
Db 961 GACATGTAGCGAATTTTCTAGAGATAGATTAAGTGTGCGGAAACGCTAACACAGGTGCTG 1020
Qy 1021 CATGGCTGTGCTGACGCTGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1080
Db 1021 CATGGCTGTGCTGACGCTGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC 1080
Qy 1081 CTTGTCAATTAATGCGATCATTTTGGTGTGGGCACTTTAATGAGATGCGCGTGACAAACCG 1140
Db 1081 CTTGTCAATTAATGCGATCATTTTGGTGTGGGCACTTTAATGAGATGCGCGTGACAAACCG 1140
Qy 1141 GAGGAAAGGTGGGAGATGAGCTGCAAGTCTCATGAGCCCTTATGGGTAGGCTTCAACGTA 1200
Db 1141 GAGGAAAGGTGGGAGATGAGCTGCAAGTCTCATGAGCCCTTATGGGTAGGCTTCAACGTA 1200
Qy 1201 TACAATGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260
Db 1201 TACAATGCGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGCGCG 1260
Qy 1261 TCGTAGTCCGATCGGAATCTGCAACTGACTCCGTGAAGTCGGAATGCTAATGATGCG 1320
Db 1261 TCGTAGTCCGATCGGAATCTGCAACTGACTCCGTGAAGTCGGAATGCTAATGATGCG 1320
Qy 1321 GGATCAGCATGTCGCGGTGAATACGTTCCGGGCTTGTACACACCGCCGCTCACACCAT 1380
Db 1321 GGATCAGCATGTCGCGGTGAATACGTTCCGGGCTTGTACACACCGCCGCTCACACCAT 1380
Qy 1381 GGGAGTGGGTTTCAACAGAGCAAGTAGTCTAACCGTAAGAGAGGCGGCTTCCACGCTGA 1440
Db 1381 GGGAGTGGGTTTCAACAGAGCAAGTAGTCTAACCGTAAGAGAGGCGGCTTCCACGCTGA 1440
Qy 1441 GATTTCATGACTGGGGTG 1457
Db 1441 GATTTCATGACTGGGGTG 1457

RESULT 6
US-10-659-983A-2
; Sequence 2, Application US/10659983A
; Publication No. US20040157313A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneq, Timothy A
; TITLE OR INVENTION: Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284779
; CURRENT APPLICATION NUMBER: US/10/659, 983A
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
```

```
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type A1 R7clone187 16S rDNA
US-10-659-983A-2

Query Match 99.7%; Score 1452.2; DB 7; Length 1457;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCCGAACGGCGACGCGATGCTTGAT 60
Db 1 ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCCGAACGGCGACGCGATGCTTGAT 60
Qy 61 CTGGTGGCGAGTGGCGGACGGGTGATGATGATGATGATGATGATGATGATGATGATGAT 120
Db 61 CTGGTGGCGAGTGGCGGACGGGTGATGATGATGATGATGATGATGATGATGATGATGAT 120
Qy 121 ACGCATGAAAGATGTGCTAATACCGCATATCTTAAAGAGAAAGACGGGATTCGAAA 180
Db 121 ACGCATGAAAGATGTGCTAATACCGCATATCTTAAAGAGAAAGACGGGATTCGAAA 180
Qy 181 GACCTGGCTTTTGGAGCGCGCATGCTGATTAAGTGTGAGTGGGTTAAAGGCTTAC 240
Db 181 GACCTGGCTTTTGGAGCGCGCATGCTGATTAAGTGTGAGTGGGTTAAAGGCTTAC 240
Qy 241 CAAGGCGACGATCAGTATGCTGAGAGGACGACGACGACGACGACGACGACGACGACG 300
Db 241 CAAGGCGACGATCAGTATGCTGAGAGGACGACGACGACGACGACGACGACGACGACG 300
Qy 301 GCCCAGACTCCTACGGAAGGACGAGTGGGAAATTTTGCAATAGGGCGCAAGCTGATC 360
Db 301 GCCCAGACTCCTACGGAAGGACGAGTGGGAAATTTTGCAATAGGGCGCAAGCTGATC 360
Qy 361 CAGCAATGCGCGGTGATGAAAGAGGCTTCCGGTTTAAAGCTTTCAGTGGAGAGA 420
Db 361 CAGCAATGCGCGGTGATGAAAGAGGCTTCCGGTTTAAAGCTTTCAGTGGAGAGA 420
Qy 421 AAAGGTTACGGTAAATATCGTACTGATGACGCTATCGACAGAGAGACACCGGCTTAC 480
Db 421 AAAGGTTACGGTAAATATCGTACTGATGACGCTATCGACAGAGAGAGACACCGGCTTAC 480
Qy 481 TACGTGCAAGCAGCGCGGTAAATCGTAGGGTGAAGCGTTAATCGAAATTAATCGGGCGT 540
Db 481 TACGTGCAAGCAGCGCGGTAAATCGTAGGGTGAAGCGTTAATCGAAATTAATCGGGCGT 540
Qy 541 AAAGGTTGCGAGCGGCTTTGTAAATGATGCAATGTAATCCCGGGCTTAACTGGGAAT 600
Db 541 AAAGGTTGCGAGCGGCTTTGTAAATGATGCAATGTAATCCCGGGCTTAACTGGGAAT 600
Qy 601 TGCCTTGAAGCTCAAGAGCTAGAGTGGGAGAGGAGTGAATTCATGTTAGCAG 660
Db 601 TGCCTTGAAGCTCAAGAGCTAGAGTGGGAGAGGAGTGAATTCATGTTAGCAG 660
Qy 661 TGAATGCGTAGAGATATGAAAGAAATCATGATGCGAAGGACCTCTCGGTTAACT 720
Db 661 TGAATGCGTAGAGATATGAAAGAAATCATGATGCGAAGGACCTCTCGGTTAACT 720
Qy 721 GACGCTCATGACCAAGCAACCGTGGGAGCAAGATTAAGATCCCTGTAGTCCAGGCC 780
Db 721 GACGCTCATGACCAAGCAACCGTGGGAGCAAGATTAAGATCCCTGTAGTCCAGGCC 780
Qy 781 CTTAAACGATGTCAACTAGTGTGGGCTTTATAGGCTTGGTAAAGCAAGCTAACGCGTGA 840
Db 781 CTTAAACGATGTCAACTAGTGTGGGCTTTATAGGCTTGGTAAAGCAAGCTAACGCGTGA 840
```

Db 781 CTAACGATGCACTAGTGTGGGCTTATTAAGCTTGGTAACGAAGCTAACCGCTGA 840  
Qy 841 AGTTGACCGCTGGGAGATACGCTGCAAGATTAATACTCAAGAAATGACGGGACCC 900  
Db 841 AGTTGACCGCTGGGAGATACGCTGCAAGATTAATACTCAAGAAATGACGGGACCC 900  
Qy 901 GCACAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTT 960  
Db 901 GCACAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTT 960  
Qy 961 GACATGTAGCCAAATTTCTAGAGATAGATTAGTCTGGGAAACGTAACAAGGCTG 1020  
Db 961 GACATGTAGCCAAATTTCTAGAGATAGATTAGTCTGGGAAACGTAACAAGGCTG 1020  
Qy 1021 CATGGCTGCTGACGCTGCTGAGATGTTGGGTTAAGTCCGCAACGACGCAAC 1080  
Db 1021 CATGGCTGCTGACGCTGCTGAGATGTTGGGTTAAGTCCGCAACGACGCAAC 1080  
Qy 1081 CTGTGATTAATTCGATTAATTCGATTAATTCGATTAATTCGATTAATTCGAT 1140  
Db 1081 CTGTGATTAATTCGATTAATTCGATTAATTCGATTAATTCGATTAATTCGAT 1140  
Qy 1141 GAGGAAAGTGGGAGATGACGCTCAAGTCTCATGCTGCTTATGGGTAGGCTTCA 1200  
Db 1141 GAGGAAAGTGGGAGATGACGCTCAAGTCTCATGCTGCTTATGGGTAGGCTTCA 1200  
Qy 1201 TACCAATGGCGGATGACAGAGGTTGCAACCCGCAAGGAGGAGCTTAATCTCA 1260  
Db 1201 TACCAATGGCGGATGACAGAGGTTGCAACCCGCAAGGAGGAGCTTAATCTCA 1260  
Qy 1261 TCGTAGTCCGATCGGATCTGCACTGCACTCGTGAAGTCCGTAATTCGATTA 1320  
Db 1261 TCGTAGTCCGATCGGATCTGCACTGCACTCGTGAAGTCCGTAATTCGATTA 1320  
Qy 1321 GGATGACGATGTCGCGGTGATTCGTTCCCGGCTTGTATACACCGCCCGTCA 1380  
Db 1321 GGATGACGATGTCGCGGTGATTCGTTCCCGGCTTGTATACACCGCCCGTCA 1380  
Qy 1381 GGGAGTGGGTTTCAACGAGAGGATGATCTAACCGTAAGGAGGCGCTTGCA 1440  
Db 1381 GGGAGTGGGTTTCAACGAGAGGATGATCTAACCGTAAGGAGGCGCTTGCA 1440  
Qy 1441 GATTGATGACTGGGCTG 1457  
Db 1441 GATTGATGACTGGGCTG 1457  
RESULT 7  
US-10-659-948A-20  
; Sequence 20, Application US/10659948A  
; Publication No. US20040101946A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovane, Timothy A  
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-294309  
; CURRENT APPLICATION NUMBER: US/10/659,948A  
; CURRENT FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573,684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386,217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,219  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 20  
; LENGTH: 1491  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: N. aestuarii-like AOB B161c1one57 16S rDNA

US-10-659-948A-20  
Query Match 95.5%; Score 1391.6; DB 7; Length 1491;  
Best Local Similarity 97.6%; Pred. No. 0; Mismatches 34; Indels 1; Gaps 1;  
Matches 1423; Conservative 0;  
Qy 1 ATTGAAAGCTGGCGCATGCTTTACATGCAAGTCGACCGGACGACGATGCTTGCAT 60  
Db 18 ATTGAAAGCTGGCGCATGCTTTACATGCAAGTCGACCGGACGACGATGCTTGCAT 77  
Qy 61 CTGTGGGAGTGGCGGACGGGTGATATGATGATGATGATGATGATGATGATGAT 120  
Db 78 CTGTGGGAGTGGCGGACGGGTGATATGATGATGATGATGATGATGATGATGAT 137  
Qy 121 ACGCATCGAAGATGCTTAATACCGCATATCTTAAGGAGGAAAGGAGGAGATCGAA 180  
Db 138 ACGCATCGAAGATGCTTAATACCGCATATCTTAAGGAGGAAAGGAGGAGATCGAA 197  
Qy 181 GACCTTGGGCTTTTGAAGCGGCGGATGCTGATTAAGTGTGTTGGGTAAAGGCTTAC 240  
Db 198 GACCTTGGGCTTTTGAAGCGGCGGATGCTGATTAAGTGTGTTGGGTAAAGGCTTAC 257  
Qy 241 CAAGCGACGATCAGTAGTGTGCTGAGAGGACGACGACCTGGGATCGAAGCAG 300  
Db 258 CAAGCGACGATCAGTAGTGTGCTGAGAGGACGACGACCTGGGATCGAAGCAG 317  
Qy 301 GCCGACATCTCTACGGGAGGACGATGGGGAATTTGGCAATGGGGCGCAAGCTGATC 360  
Db 318 GCCGACATCTCTACGGGAGGACGATGGGGAATTTGGCAATGGGGCGCAAGCTGATC 377  
Qy 361 CAGCAATGGCGGTGATGTAAGAAAGGCTTGGGTTGTAAGCTCTTTCAGTGAAGAA 420  
Db 378 CAGCAATGGCGGTGATGTAAGAAAGGCTTGGGTTGTAAGCTCTTTCAGTGAAGAA 437  
Qy 421 AAAGTTACGATTAATATCGTACTCATGACGATTCGACAGAAAGACACCGGCTAAC 480  
Db 438 AAAGTTACGATTAATATCGTACTCATGACGATTCGACAGAAAGACACCGGCTAAC 497  
Qy 481 TAGTGCACGACGCGCGGTAAATGATGAGGTGCAAGCCTTAATCGGAATTCCTGG 540  
Db 498 TAGTGCACGACGCGCGGTAAATGATGAGGTGCAAGCCTTAATCGGAATTCCTGG 557  
Qy 541 AAAGGTGGCGGCGGCTTGTAAAGTCAATGTAATTCGCGGCTTAACTCGGGAAT 600  
Db 558 AAAGGTGGCGGCGGCTTGTAAAGTCAATGTAATTCGCGGCTTAACTCGGGAAT 617  
Qy 601 TGGCTTTGAAACTCAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGACAG 660  
Db 618 TGGCTTTGAAACTCAAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGACAG 677  
Qy 661 TGAATGCGTAGATATGAGAAACATGATGCGAAGGCGACCTCTCGGTTAACTCT 720  
Db 678 TGAATGCGTAGATATGAGAAACATGATGCGAAGGCGACCTCTCGGTTAACTCT 737  
Qy 721 GACGCTCATGACGAAAGGTGGGAGCAAAAGATTAATGATACCTCGTGTGACAGCC 780  
Db 738 GACGCTCATGACGAAAGGTGGGAGCAAAAGATTAATGATACCTCGTGTGACAGCC 797  
Qy 781 CTAACGATGCACTAGTGTGTTGGCCTTAATTAAGCTTGTGAACGAGCTTAACGCGTGA 840  
Db 798 CTAACGATGCACTAGTGTGTTGGCCTTAATTAAGCTTGTGAACGAGCTTAACGCGTGA 857  
Qy 841 AGTTGACCGCTGGGAGATACGCTGCAAGATTAATACTCAAGAAATGACGGGACCC 900  
Db 858 AGTTGACCGCTGGGAGATACGCTGCAAGATTAATACTCAAGAAATGACGGGACCC 917  
Qy 901 GCACAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTT 960  
Db 918 GCACAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTT 977  
Qy 961 GACATGTAGCCAAATTTCTAGAGATAGATTAGTCTGGGAAACGTAACAAGGCTG 1019  
Db 978 GACATGTAGCCAAATTTCTAGAGATAGATTAGTCTGGGAAACGTAACAAGGCTG 1037

QY 1020 GCATGGCTGTCTCAGCTCTGTCTCGTAGATTTGGTTAACTCCCGCAACGAGCGCAAC 1079  
DB 1038 GCATGGCTGTCTCAGCTCTGTCTCGTAGATTTGGTTAACTCCCGCAACGAGCGCAAC 1097  
QY 1080 CCTTGTCAATTAATTCATTCATTTGGTGGCACTTTAATGAGACTGCGCGTAGCAAAAC 1139  
DB 1098 CTTTGTCAATTAATTCATTCATTTAGTTGGCACTTTAATGAGACTGCGCGTAGCAAAAC 1157  
QY 1140 GGAGGAGGTGGGGATGAGCTCAAGTCTCATGCCCCCTTATGGGTAGGGCTTCAACGTA 1199  
DB 1158 GGAGGAGGTGGGGATGAGCTCAAGTCTCATGCCCCCTTATGGGTAGGGCTTCAACGTA 1217  
QY 1200 ATACAAATGCGCGTCAAGAGGGTTGCCAACCCGCGAGGGGGAGCTAATCTCGAAGAGGC 1259  
DB 1218 ATACAAATGCGCGTCAAGAGGGTTGCCAACCCGCGAGGGGGAGCTAATCTCGAAGAGGC 1277  
QY 1260 GTCTGAGTCCGGATCGAGTGTGCAACTGCACTCCGTGAAATCGGAAATCGTAGTAATCG 1319  
DB 1278 GTCTGAGTCCGGATCGAGTGTGCAACTGCACTCCGTGAAATCGGAAATCGTAGTAATCG 1337  
QY 1320 CGGATCAAGATGCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCACACCA 1379  
DB 1338 CGGATCAAGATGCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCACACCA 1397  
QY 1380 TGGGATGGGTTTACCAAGAGCAGTACTTAACCGTGAAGAGGGCGCTTGGCCACGGTG 1439  
DB 1398 TGGGATGGGTTTACCAAGAGCAGTACTTAACCGTGAAGAGGGCGCTTGGCCACGGTG 1457  
QY 1440 AGATTCATGACTGGGGTG 1457  
DB 1458 AGATTCATGACTGGGGTG 1475

RESULT 8  
US-10-659-980A-20  
Sequence 20, Application US/10659980A  
Publication No. US20040106133A1  
GENERAL INFORMATION:  
APPLICANT: Hovaneec, Timothy A  
TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria  
FILE REFERENCE: 81289-284781  
CURRENT APPLICATION NUMBER: US/10/659,980A  
CURRENT FILING DATE: 2003-09-10  
PRIOR APPLICATION NUMBER: US 09/573,684  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 60/386,217  
PRIOR FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: US 60/386,218  
PRIOR FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: US 60/386,219  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 20  
LENGTH: 1491  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: N. Aestuarii-like AOB B16clone57 16S rDNA  
US-10-659-980A-20

Query Match 95.5%; Score 1391.6; DB 7; Length 1491;  
Best Local Similarity 97.6%; Pred. No. 0;  
Matches 1423; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 1 ATTTGAACCTGCGCGCATCTTTACATGCAAGTCGAAACGCGACACGAGATGCTTGAT 60  
DB 18 ATTGAACCTGCGCGCATCTTTACATGCAAGTCGAAACGCGACACGAGATGCTTGAT 77  
QY 61 CTGCTGAGAGTGGGCGAGCGGCTGATATGATCGAAGCGATGACAGAGAGGCGGTA 120  
DB 78 CTGCTGAGAGTGGGCGAGCGGCTGATATGATCGAAGCGATGACAGAGAGGCGGTA 137

QY 121 ACGCATGGAAGATGTGCTAATACCGCATATACTCTAAGAGGAAGAGCGGAGATCGAAA 180  
DB 138 ACGCATGGAAGATGTGCTAATACCGCATATACTCTAAGAGGAAGAGCGGAGATCGAAA 197  
QY 181 GACCTTGCGCTTTTGAACCGCGCATGCTCTGATTAAGTAGTTGGTGGGTAAGCGCTTAC 240  
DB 198 GACCTTGCGCTTTTGAACCGCGCATGCTCTGATTAAGTAGTTGGTGGGTAAGCGCTTAC 257  
QY 241 CAAGCGACATCACTAGTATGCTCTGAGAGGACCAACCGACCACTGGGACTGAGACACG 300  
DB 258 CAAGCGACATCACTAGTATGCTCTGAGAGGACCAACCGACCACTGGGACTGAGACACG 317  
QY 301 GCCCAACTCCTTACGGGAGGAGCAGATGGGGGAATTTTGAACATAGGCGCAAGCCTGATC 360  
DB 318 GCCCAACTCCTTACGGGAGGAGCAGATGGGGGAATTTTGAACATAGGCGCAAGCCTGATC 377  
QY 361 CAGCAATGCCCGTGAAGTGAAGAGGCGCTTCGGGTTGTAAGCTCTTCACTCGAAGAA 420  
DB 378 CAGCAATGCCCGTGAAGTGAAGAGGCGCTTCGGGTTGTAAGCTCTTCACTCGAAGAA 437  
QY 421 AAAGTTACGGTAAATTAATCTGACTCATGACGGTATGACAGAAAGCAACCGGCTAAC 480  
DB 438 AAAGTTACGGTAAATTAATCTGACTCATGACGGTATGACAGAAAGCAACCGGCTAAC 497  
QY 481 TACGTCCAGCAGCGCGGTAAATCGTAGGGTGAAGGGGTTAAATCGGAATTAATGAGGCGT 540  
DB 498 TACGTCCAGCAGCGCGGTAAATCGTAGGGTGAAGGGGTTAAATCGGAATTAATGAGGCGT 557  
QY 541 AAAGGTGCGCAGCGCGCTTGTGAAGTCAATGTGAATCCCGGGCTTAACTGGGAAT 600  
DB 558 AAAGGTGCGCAGCGCGCTTGTGAAGTCAATGTGAATCCCGGGCTTAACTGGGAAT 617  
QY 601 TGCCTTGAATCTACAAAGCTTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTAGAG 660  
DB 618 TGCCTTGAATCTACAAAGCTTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTAGAG 677  
QY 661 TGAATGCGTGAAGTATGGAAGAACATGATGCGAAGGAGCGCTCGTGGTTAACT 720  
DB 678 TGAATGCGTGAAGTATGGAAGAACATGATGCGAAGGAGCGCTCGTGGTTAACT 737  
QY 721 GACCTCATGCAAGAAACGCTGGGAGGCAACAGAGATTAGTACCTGTAGTCCAGCC 780  
DB 738 GACCTCATGCAAGAAACGCTGGGAGGCAACAGAGATTAGTACCTGTAGTCCAGCC 797  
QY 781 CTAAACGATGCACTAGTGTGGGCTTATTTAGGCTTGTGAACGAAGCTAACGCGTGA 840  
DB 798 CTAAACGATGCACTAGTGTGGGCTTATTTAGGCTTGTGAACGAAGCTAACGCGTGA 857  
QY 841 AGTTGACCGCTGGGGAGTACGCTCGCAAGATTAACTCAAGAGAAATTGACGGGGACCC 900  
DB 858 AGTTGACCGCTGGGGAGTACGCTCGCAAGATTAACTCAAGAGAAATTGACGGGGACCC 917  
QY 901 GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAGAAACCTTACCTT 960  
DB 918 GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGCGAAGAAACCTTACCTT 977  
QY 961 GACATGTAAGCAATTTTCTAGAGATGATTAATGCTGCTTGGGAAACGCTAACACAGTGTCT 1019  
DB 978 GACATGTAAGCAATTTTCTAGAGATGATTAATGCTGCTTGGGAAACGCTAACACAGTGTCT 1037  
QY 1020 GCATGCGTGTGTGCACTCGCTGTGTGAGAGTTGGGTTAATGCTCCGCAACGAGCGCAAC 1079  
DB 1038 GCATGCGTGTGTGCACTCGCTGTGTGAGAGTTGGGTTAATGCTCCGCAACGAGCGCAAC 1097  
QY 1080 CTTTGTCAATTAATTCATTCATTTGGTGGCACTTTAATGAGACTGCGCGTAGCAAAAC 1139  
DB 1098 CTTTGTCAATTAATTCATTCATTTAGTTGGCACTTTAATGAGACTGCGCGTAGCAAAAC 1157  
QY 1140 GGAGGAGGTGGGATGAGCTCAAGTCTCATGCCCCCTTATGGGTAGGGCTTCAACGTA 1199  
DB 1158 GGAGGAGGTGGGATGAGCTCAAGTCTCATGCCCCCTTATGGGTAGGGCTTCAACGTA 1217

QY 1200 ATACATGCGCGTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAAGAAAGCC 1259  
DB 1218 ATACATGCGCGTACAGAGGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAAGAAAGCC 1277  
QY 1260 GTCTGATGCGGATCGGAGTCTGCAACTCGACTCCGTGAAAGCGGAATCGAGTAATCG 1319  
DB 1278 GTCTGATGCGGATCGGAGTCTGCAACTCGACTCCGTGAAAGCGGAATCGAGTAATCG 1337  
QY 1320 CGGATCAGATGTCGCGGTGAATACGTTCCCGGGTCTTGTAACAACCGCCGTCACACA 1379  
DB 1338 CGGATCAGATGTCGCGGTGAATACGTTCCCGGGTCTTGTAACAACCGCCGTCACACA 1397  
QY 1380 TGGAGTGGGTTTCAACGAAGCAGGTACTTAACCGTAAAGAGGGCCCTTGCCACGGTG 1439  
DB 1398 TGGAGTGGGTTTCAACGAAGCAGGTACTTAACCGTAAAGAGGGCCCTTGCCACGGTG 1457  
QY 1440 AGATTCAATGCTGGGGGTG 1457  
DB 1458 AGATTCAATGCTGGGGGTG 1475

## RESULT 9

US-10-659-983A-20  
; Sequence 20, Application US/10659983A  
; Publication No. US20040157313A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovaneq, Timothy A  
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-284779  
; CURRENT APPLICATION NUMBER: US/10/659,983A  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573,684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386,217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,219  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 20  
; LENGTH: 1491  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: N. Aestuarii-like MOB B16clones7 16S rDNA  
US-10-659-983A-20

Query Match 95.5%; Score 1391.6; DB 7; Length 1491;  
Best Local Similarity 97.6%; Pred. No. 0;  
Matches 1433; Conservative 0; Mismatches 34; Indels 1; Gaps 1;  
QY 1 ATTTGAACGCTGGCGGATGCTTTTACACATGCAAGTGAACGGGAGACCGAGTCTTGCAAT 60  
DB 18 ATTTGAACGCTGGCGGATGCTTTTACACATGCAAGTGAACGGGAGACCGAGTCTTGCAAT 77  
QY 61 CTGTGTCGCAATGTCGCGGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 120  
DB 78 CTGTGTCGCAATGTCGCGGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 137  
QY 121 AGGCATCGAAAGATGCTTAATACCGCATATCTTAAGAGGAAAGCAGGGATCGAAA 180  
DB 138 AGGCATCGAAAGATGCTTAATACCGCATATCTTAAGAGGAAAGCAGGGATCGAAA 197  
QY 181 GACCTTGCGCTTTTGAAGCGCGCATGCTGATTAAGTAAAGGCTTAAAGGCTTAAAGGCTTAA 240  
DB 198 GACCTTGCGCTTTTGAAGCGCGCATGCTGATTAAGTAAAGGCTTAAAGGCTTAAAGGCTTAA 257  
QY 241 CAAGGCGCATGATGATGTTGTTCTGAGAGAGACGACCAAGCAGCTGAGACAG 300  
DB 258 CAAGGCGCATGATGATGTTGTTCTGAGAGAGACGACCAAGCAGCTGAGACAG 317

QY 301 GCCAGACTCCTACGGGAGCAGCAGTGGGAAATTTTGAACATATGGCGCAAGCCTGATC 360  
DB 318 GCCAGACTCCTACGGGAGCAGCAGTGGGAAATTTTGAACATATGGCGCAAGCCTGATC 377  
QY 361 CAGCAATGCGCGGTGAGTGAAGAAAGCCTTGGGTTGTAAGCTCTTTCACTCGAAGAA 420  
DB 378 CAGCAATGCGCGGTGAGTGAAGAAAGCCTTGGGTTGTAAGCTCTTTCACTCGAAGAA 437  
QY 421 AAAGTTACGGTAAATTAATCTGACTCATGACCGTATGACAGAAAGCAACCGGCTAAC 480  
DB 438 AAAGTTACGGTAAATTAATCAACTTAATGACCGTATGACAGAAAGCAACCGGCTAAC 497  
QY 461 TACGTGCAAGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAATCTGGGGGT 540  
DB 498 TACGTGCAAGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAATCTGGGGGT 557  
QY 541 AAAGGTCGCGAGCGCGCTTTGTAAGTGAATGTAATCCCGGGCTTAACTTGGGAAT 600  
DB 558 AAAGGTCGCGAGCGCGCTTTGTAAGTGAATGTAATCCCGGGCTTAACTTGGGAAT 617  
QY 601 TCGGTTGAAACTACAGAGCTAGAGTGGCGAGAGAGTGGAAATTCATGTTGTAAGAG 660  
DB 618 TCGGTTGAAACTACAGAGCTAGAGTGGCGAGAGAGTGGAAATTCATGTTGTAAGAG 677  
QY 661 TGAATTGGGTAGATTAATGAAAGAACATCGAATGGGAAAGCGCCTCGGTTAACT 720  
DB 678 TGAATTGGGTAGATTAATGAAAGAACATCGAATGGGAAAGCGCCTCGGTTAACT 737  
QY 721 GACGCTCATGACGAAGACGCTGGGAGCAACAGAGTTAGATACCTGTTAGTCCAGCGC 780  
DB 738 GACGCTCATGACGAAGACGCTGGGAGCAACAGAGTTAGATACCTGTTAGTCCAGCGC 797  
QY 781 CTAAACGATGCTACTAGTTGTTGGGCTTTATTAAGCTTGGTAAGAACTAACCGGTGA 840  
DB 798 CTAAACGATGCTACTAGTTGTTGGGCTTTATTAAGCTTGGTAAGAACTAACCGGTGA 857  
QY 841 AGTTGACCGCTCGGGAAGTACGCTGCAAGATTAATACTCAAGGAATGACGGGAGCC 900  
DB 858 AGTTGACCGCTCGGGAAGTACGCTGCAAGATTAATACTCAAGGAATGACGGGAGCC 917  
QY 901 GCACAAGCGGTGATTAATGATTAATTCATGCAACCGCAAAAACCTTACCTACCTT 960  
DB 918 GCACAAGCGGTGATTAATGATTAATTCATGCAACCGCAAAAACCTTACCTACCTT 977  
QY 961 GACATGATGCAATTTTCTAGATATGATTAGT-CTTGGGAGACGCTTAACACAGTGTCT 1019  
DB 978 GACATGATGCAATTTTCTAGATATGATTAGT-CTTGGGAGACGCTTAACACAGTGTCT 1037  
QY 1020 GCATGCTGTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079  
DB 1038 GCATGCTGTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1097  
QY 1080 CCTGTGATTAATGTCATCATTTGGTGGGACCTTAATGAGACTGCGGTGACAAAC 1139  
DB 1098 CCTGTGATTAATGTCATCATTTGGTGGGACCTTAATGAGACTGCGGTGACAAAC 1157  
QY 1140 GAGGAGAGGTGGGATGAGCTGCAAGTCTCATAGCCCTTAATGAGGCTTCAACGTA 1199  
DB 1158 GAGGAGAGGTGGGATGAGCTGCAAGTCTCATAGCCCTTAATGAGGCTTCAACGTA 1217  
QY 1200 ATACATGCGCGTACAGAGGGTTGCAACCGCGAGGGGAGCTTAATCTCAAGAAAGCGC 1259  
DB 1218 ATACATGCGCGTACAGAGGGTTGCAACCGCGAGGGGAGCTTAATCTCAAGAAAGCGC 1277  
QY 1260 GTCTGATGCGGATCGGAGTCTGCAACTCGACTCCGTGAAAGCGGAATCGAGTAATCG 1319  
DB 1278 GTCTGATGCGGATCGGAGTCTGCAACTCGACTCCGTGAAAGCGGAATCGAGTAATCG 1337  
QY 1320 CGGATCAGATGTCGCGGTGAATACGTTCCCGGGTCTTGTAACAACCGCCGTCACACA 1379  
DB 1338 CGGATCAGATGTCGCGGTGAATACGTTCCCGGGTCTTGTAACAACCGCCGTCACACA 1397  
QY 1380 TGGAGTGGGTTTCAACGAAGCAGGTACTTAACCGTAAAGAGGGCGCTTGCCACGGTG 1439

Db 1398 TGGAGTGGCTTTCACCGAAGCAGATGCTAACCGTAAGAGGCGCTTGGCCACGGTG 1457  
Qy 1440 AGATTCATGACTGGGGTG 1457  
Db 1458 AGATTCATGACTGGGGTG 1475  
RESULT 10  
US-10-659-948A-19  
; Sequence 19, Application US/10659948A  
; Publication No. US20040101946A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovonec, Timothy A  
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-294309  
; CURRENT APPLICATION NUMBER: US/10/659,948A  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573,684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386,217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386,219  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 19  
; LENGTH: 1494  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: N. Aestuarii-like AOB P4clone31 16S rDNA  
US-10-659-948A-19  
Query Match 93.3% Score 1359.2; DB 7; Length 1494;  
Best Local Similarity 96.3%; Pred. No. 0;  
Matches 1402; Conservative 0; Mismatches 53; Indels 1; Gaps 1;  
Qy 1 ATTGAACGCTGCGGCATGCTTTTACATGCAATGCAAGCGCAGACGATGCTTGCAT 60  
Db 19 ATTGAACGCTGCGGCATGCTTTTACATGCAATGCAAGCGCAGACGATGCTTGCAT 78  
Qy 61 CTGTGGCGAGTGGCGGACGGGTGATATGATGATGCAATGATATCAAGAAAGGGGGTA 120  
Db 79 CTGTGGCGAGTGGCGGACGGGTGATATGATGATGCAATGATATCAAGAAAGGGGGTA 138  
Qy 121 ACGCATGGAAGATGCTAATACCGCATATCTCTAAGAGGAAGAGGAGATCGAAA 180  
Db 139 ACGCATGGAAGATGCTAATACCGCATATCTCTAAGAGGAAGAGGAGATCGAAA 198  
Qy 181 GACCTTGCGCTTTTGAAGCGGCCGATGTCGATTAAGTATGTTGGGTAAAGGCTTAC 240  
Db 199 GACCTTGCGCTTTTGAAGCGGCCGATGTCGATTAAGTATGTTGGGTAAAGGCTTAC 258  
Qy 241 CAAGCGCAGATCAATGATGTTGGTCTGAAGAGACACGACCACTGGGACTGAAGACG 300  
Db 259 CAAGCGCAGATCAATGATGTTGGTCTGAAGAGACACGACCACTGGGACTGAAGACG 318  
Qy 301 GCCGAGCTCTTAAGGAGGAGCAGATGGGGATTTTGAACATGAGGCGCAAGCTGTATC 360  
Db 319 GCCGAGCTCTTAAGGAGGAGCAGATGGGGATTTTGAACATGAGGCGCAAGCTGTATC 378  
Qy 361 CAGCAATGCCGCTGATGGAAGAGCCCTTCGGGTGTAAAGCTTTTCACTGAGAAAG 420  
Db 379 CAGCAATGCCGCTGATGGAAGAGCCCTTCGGGTGTAAAGCTTTTCACTGAGAAAG 438  
Qy 421 AAAGGTTACGTTAAATATCTGACTCATGACGGTATGACAGAAAGACCGGCTTAC 480  
Db 439 AAAGGTTACGTTAAATATCTGACTCATGACGGTATGACAGAAAGACCGGCTTAC 498  
Qy 481 TAGCTGCAAGACGCGCGGTAAATACGTAGGGTGAAGGTTAAATGCAATTAATCTGGGCT 540

Db 499 TACGTGCAGACGCGCGGTAAATACGTAGGGTGAAGCGTTAAATCGAAATTAATCTGGGCT 558  
Qy 541 AAAGGTTACGCGCGGCTTTGTAAGTCAATGTAATATCCCGGGCTTAACTGGGAAT 600  
Db 559 AAAGGTTACGCGCGGCTTTGTAAGTCAATGTAATATCCCGGGCTTAACTGGGAAT 618  
Qy 601 TGCCTTTGAATCTCAAGAGCTAGAGTGGCAGAGGAGGTGAATTCATGTTAGCAG 660  
Db 619 TGCCTTTGAATCTCAAGAGCTAGAGTGGCAGAGGAGGTGAATTCATGTTAGCAG 678  
Qy 661 TGAATGCGTAGAGATATGGAAGAACATCGATGCGAAGGCGAGCTTCTGGTTAACT 720  
Db 679 TGAATGCGTAGAGATATGGAAGAACATCGATGCGAAGGCGAGCTTCTGGTTAACT 738  
Qy 721 GACGCTCATGACGAAAGCGTGGGAGCAACACAGATTAATGATACCTGTTGTCCAGCC 780  
Db 739 GACGCTCATGACGAAAGCGTGGGAGCAACACAGATTAATGATACCTGTTGTCCAGCC 798  
Qy 781 CTTAAGCATGTCATGATGTTGGCTTATTAAGCTTGGTTAAGAACTAAGCCTAAGCCTGA 840  
Db 799 CTTAAGCATGTCATGATGTTGGCTTATTAAGCTTGGTTAAGAACTAAGCCTAAGCCTGA 858  
Qy 841 AGTTGACCGCTGGGGAGTACGTTGCGAAGATTAAACTCAAGGAATTGACGGGACCC 900  
Db 859 AGTTGACCGCTGGGGAGTACGTTGCGAAGATTAAACTCAAGGAATTGACGGGACCC 918  
Qy 901 GCACAAAGCGTGAATTTATGCGATTAATGCAATGCAACGCGAATACTTACCTTACCCTT 960  
Db 919 GCACAAAGCGTGAATTTATGCGATTAATGCAATGCAACGCGAATACTTACCTTACCCTT 978  
Qy 961 GACATGTAGCAATTTTCTAGAGATGATTAAGTGTGTTGGGAGTAAAGCTTACACAGTGTCT 1019  
Db 979 GACATGTAGCAATTTTCTAGAGATGATTAAGTGTGTTGGGAGTAAAGCTTACACAGTGTCT 1038  
Qy 1020 GCATGCTGTGTCAGCTGCTGTCGTGAGATGTTGGTTAAAGTCCCGCAACGACGCAAC 1079  
Db 1039 GCATGCTGTGTCAGCTGCTGTCGTGAGATGTTGGTTAAAGTCCCGCAACGACGCAAC 1098  
Qy 1080 CTTGTGATTAATGTCATCATCTTTGGTTGGGCACTTTAATGAGCTGCGGTGAACAAACC 1139  
Db 1099 CTTGTGATTAATGTCATCATCTTTGGTTGGGCACTTTAATGAGCTGCGGTGAACAAACC 1158  
Qy 1140 GGAAGAAAGTGGGATGACGTCAGTCTTCAATGAGCTTCAATGAGGAGTAACTTCAAGAAAGCGC 1259  
Db 1159 GGAAGAAAGTGGGATGACGTCAGTCTTCAATGAGCTTCAATGAGGAGTAACTTCAAGAAAGCGC 1278  
Qy 1200 ATACATGCGCGCTGACAGAGGCTTGCACCCCGAGGAGGAGTAACTTCAAGAAAGCGC 1278  
Db 1219 ATACATGCGCGCTGACAGAGGCTTGCACCCCGAGGAGGAGTAACTTCAAGAAAGCGC 1278  
Qy 1260 GTGCTGATGCGGATGCGAGTCTGCACTGCACTCCGTAAGTCCGTAATCGTAAATTCG 1319  
Db 1279 GTGCTGATGCGGATGCGAGTCTGCACTGCACTCCGTAAGTCCGTAATCGTAAATTCG 1338  
Qy 1320 CGGATGACGATGTCGCGGTGAATGCTTCCGGGCTTGTACACACGCGCTTACACCA 1379  
Db 1339 CGGATGACGATGTCGCGGTGAATGCTTCCGGGCTTGTACACACGCGCTTACACCA 1398  
Qy 1380 TGGAGTGGCTTTTCAAGAAAGCAGTATGTTAACTGTAAGAGAGGCGCTTGGCAAGGCTG 1439  
Db 1399 TGGAGTGGCTTTTCAAGAAAGCAGTATGTTAACTGTAAGAGAGGCGCTTGGCAAGGCTG 1458  
Qy 1440 AGATTCATGACTGGGG 1455  
Db 1459 TGTCAATGACTTGGG 1474  
RESULT 11  
US-10-659-980A-19  
; Sequence 19, Application US/10659980A  
; Publication No. US20040106133A1  
; GENERAL INFORMATION:

```
APPLICANT: Hovanec, Timothy A
TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria
FILE REFERENCE: 81289-284781
CURRENT APPLICATION NUMBER: US/10/659,980A
CURRENT FILING DATE: 2003-09-10
PRIOR APPLICATION NUMBER: US 09/573,684
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 60/386,217
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: US 60/386,218
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: US 60/386,219
PRIOR FILING DATE: 2002-09-19
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 19
LENGTH: 1494
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: N. Aestuarii-like MOB pAclone31 16S rDNA
US-10-659-980A-19

Query Match      93.3%; Score 1359.2; DB 7; Length 1494;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1402; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

QY      1  ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAAACGCGACGCGATGCTTGAT 60
DB      19  ATTGAACGCTGGCGGCGATCTTTACATGCAAGTCGAAACGCGACGCGATGCTTGAT 78
QY      61  CTGTGGCGAGTGGCGGAGCGGGTGAAGTGCATCGGAAAGTATCCAGAAAGGGGGGTA 120
DB      79  CTGTGGCGAGTGGCGGAGCGGGTGAAGTGCATCGGAAAGTATCCAGAAAGGGGGGTA 138
QY      121  AGCGATCGAAAGATGTCTTAATCCGCATATCTCTTAAGAGAAAGCAGGGGATGAAA 180
DB      139  AGCGATCGAAAGATGTCTTAATCCGCATATCTCTTAAGAGAAAGCAGGGGATGAAA 198
QY      181  GACCTTGGCGCTTTGGAGGGCGGATGCTGATTAAGTGGTGGGGTAAAGGCTTAC 240
DB      199  GACCTTGGCTTTGGAGGGCGGATGCTGATTAAGTGGTGGGGTAAAGGCTTAC 258
QY      241  CAAGGCGAGATCATGATGTGTCTGAGAGACGACCAAGCCACCTGGGACTGAGACAG 300
DB      259  CAAGGCGAGATCATGATGTGTCTGAGAGAGACCAAGCCACCTGGGACTGAGACAG 318
QY      301  GCCCAGACTCTTACCGGAGGCGACAGTGGGAAATTTTGCACATGGCGCGACCTGATC 360
DB      319  GCCCAGACTCTTACCGGAGGCGACAGTGGGAAATTTTGCACATGGCGCGACCTGATC 378
QY      361  CAGCAATGCCGCGCTGATGAAAGGCTTGGGTTGTAAGCTTTTCATGAGAGAA 420
DB      379  CAGCAATGCCGCGCTGATGAAAGGCTTGGGTTGTAAGCTTTTCATGAGAGAA 438
QY      421  AAAGTTACGGTAAATATCGTACATGACCGTATCGACAGAAAGACGCGGCTAAC 480
DB      439  AAAGTTACGGTAAATATCGTACATGACCGTATCGACAGAAAGACGCGGCTAAC 498
QY      481  TACGTCCAGACGCGCGGTAATACGTAAGGTGCAAGCGTTAATCGGAATTTACTGGGCGT 540
DB      499  TACGTCCAGACGCGCGGTAATACGTAAGGTGCAAGCGTTAATCGGAATTTACTGGGCGT 558
QY      541  AAAGGTGGCAGCGCGCTTTGTAAGTCAGATGTAATCCCGGGGCTTAACTGGGAAT 600
DB      559  AAAGGTGGCAGCGCGCTTTGTAAGTCAGATGTAATCCCGGGGCTTAACTGGGAAT 618
QY      601  TGGCTTTGAATCAAGGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTAGCAG 660
DB      619  TGGCTTTGAATCAAGGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTAGCAG 678
QY      661  TGAATTCGTAGATATGGAAGAACATGCAATGGCGAAGCGACCTCTCTGGGTTACACT 720
```

```
DB      679  TGAATTCGTAGATATGGAAGAACATGCAATGGCGAAGCGACCCCTCTGGGTTACACT 738
QY      721  GACGCTATGCAACGAACCGTGGGAGCAACAGGATTAATCCCTGGTATGCCAGCC 780
DB      739  GACGCTATGCAACGAACCGTGGGAGCAACAGGATTAATCCCTGGTATGCCAGCC 798
QY      781  CTAAACGATGTCATTAATGTTGGGCTTTATTAAGCTTGGTAACGAAGCTAACCGTGA 840
DB      799  CTAAACGATGTCATTAATGTTGGGCTTTATTAAGCTTGGTAACGAAGCTAACCGTGA 858
QY      841  AGTTACCGCTGGGAGTACGCTGCAAGATTAATACTCAAGGAATTCAGCGGACCC 900
DB      859  AGTTACCGCTGGGAGTACGCTGCAAGATTAATACTCAAGGAATTCAGCGGACCC 918
QY      901  GCACAGCGGTGGATTAATGTTGATTAATTCGATGCAAGCGGAAACCTTACTCCCTT 960
DB      919  GCACAGCGGTGGATTAATGTTGATTAATTCGATGCAAGCGGAAACCTTACTCCCTT 978
QY      961  GACATGTAGCGAATTTTCTAGATAGATTAGTG-CTTCGGGAACGCTTAACAGGTGCT 1019
DB      979  GACATGTAGCGAATTTTCTAGATAGATTAGTGCTTCGGGAACGCTTAACAGGTGCT 1038
QY      1020  GCATGCTGTCTCAGCTCGTGTCTGAGATGTTGGTTAAGTCCCGCAACGAGCGAAC 1079
DB      1039  GCATGCTGTCTCAGCTCGTGTCTGAGATGTTGGTTAAGTCCCGCAACGAGCGAAC 1098
QY      1080  CCTTGTCAATTAATGCCATCATTTGGTGGGACCTTTAATGAGACTGCCGGTGAACAA 1139
DB      1099  CCTTGTCAATTAATGCCATCATTTAGTGGGACCTTTAATGAGACTGCCGGTGAACAA 1158
QY      1140  GGAGGAAGGTGGGGATGACGTAAGTCCATGAGCCCTTAATGGGTAGGGCTTCAACGTA 1199
DB      1159  GGAGGAAGGTGGGGATGACGTAAGTCCATGAGCCCTTAATGGGTAGGGCTTCAACGTA 1218
QY      1200  ATACATGCGCGGTACAGAGGCTTCCAAACCCGCAAGGGGAGCTAATTCAGAAAGCC 1259
DB      1219  ATACATGCGCGGTACAGAGGCTTCCAAACCCGCAAGGGGAGCTAATTCAGAAAGCC 1278
QY      1260  GTGTAGTCCGATGGAGTCTGCAACTGCACTCCGTGAAGTCCGTAATGGCTTAATG 1319
DB      1279  GTGTAGTCCGATGGAGTCTGCAACTGCACTCCGTGAAGTCCGTAATGGCTTAATG 1338
QY      1320  CGGATCAGATATCCCGGGTAAATAGTTTCCGGGCTTGTACACACCGCCGTCAACCA 1379
DB      1339  CGGATCAGATATCCCGGGTAAATAGTTTCCGGGCTTGTACACACCGCCGTCAACCA 1398
QY      1380  TGGAGTGGGTTTACCAAGACAGGTAAGTCTTAACCGTAAGAGGCGCTTGCACGGTG 1439
DB      1399  TGGAGTGGGTTTACCAAGACAGGTAAGTCTTAACCGTGGAGAGCGCTTACACGGTG 1458
QY      1440  AGATTCAATGACTGGGG 1455
DB      1459  TGGTCAATGACTGGGG 1474

RESULT 12
US-10-659-983A-19
Sequence 19, Application US/10659983A
Publication No. US20040157313A1
GENERAL INFORMATION:
APPLICANT: Hovanec, Timothy A
TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
FILE REFERENCE: 81289-284779
CURRENT APPLICATION NUMBER: US/10/659,983A
CURRENT FILING DATE: 2003-09-10
PRIOR APPLICATION NUMBER: US 09/573,684
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 60/386,217
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: US 60/386,218
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: US 60/386,219
```



```

; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 19
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: N. Aestuarii-like AOB Paclione31 16S rDNA
US-10-659-983A-19

Query Match      93.3%; Score 1359.2; DB 7; Length 1494;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1402; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

Qy      1  ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGGCGACACGAGATGCTTGAT 60
Db      19  ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGGCGACGAGGCTTTGAC 78
Qy      61  CTGGTGGAGAGTGGCGGACGGGTGATGATGATGCAACGATTCAGAAAGAGGGGGTA 120
Db      79  CTGGTGGAGAGTGGCGGACGGGTGATGATGATGCAACGATTCAGAAAGTGGGGATTA 138
Qy      121  ACGCATCGAAAGATGTGTCTAATACCGCATATCTTAAGAGAGAAAGCGGSAATCGAAA 180
Db      139  ACGCATCGAAAGATGTGTCTAATACCGCATATCTCTTACGAGAGAAAGCGGSAATCGAAA 198
Qy      181  GACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTAAAGGCTTAC 240
Db      199  GACCTTGCGCTTTTGGAGCGGCGCATGTCTGATTAGCTAGTTGGTGGGTAAAGGCTTAC 258
Qy      241  CAAGCGGACGATCGATGATGTGTGAGAGGACGACGACCACTGGGACTGAGACACG 300
Db      259  CAAGCGGACGATCGATGATGTGTGAGAGGACGACGACCACTGGGACTGAGACACG 318
Qy      301  GCCGAGACTCCTTACGAGAGGACGACGATGGGGAATTTTGGACATGSGCGCAAGCTGATC 360
Db      319  GCCGAGACTCCTTACGAGAGGACGACGATGGGGAATTTTGGACATGSGCGCAAGCTGATC 378
Qy      361  CAGCAATGCCCGGTGATGAGAAAGCCTTGGGGTTGTAAGCTTTTCACTCGAAGAA 420
Db      379  CAGCAATGCCCGGTGATGAGAAAGCCTTGGGGTTGTAAGCTTTTCACTCGAAGAA 438
Qy      421  AAAAGTTACGGTAAATATCGTGACTCATGACGGTATGACAGAAAGAACCGGCTTAC 480
Db      439  AAAAGTTGTGACTAAATATCAACAATTATGACGGTATGACAGAAAGAACCGGCTTAC 498
Qy      481  TACGTGCGACAGACGCGCGTAAATCGTAGGGTGCAGCGTTAATCGAATTACTGGCGGT 540
Db      499  TACGTGCGACAGACGCGCGTAAATCGTAGGGTGCAGCGTTAATCGAATTACTGGCGGT 558
Qy      541  AAAAGGTGCGAGCGCGCTTTGTAAATGATGAAATGCCCGGGCTTTAACTTGGGAAT 600
Db      559  AAAAGGTGCGAGCGCGCTTTGTAAATGATGAAATGCCCGGGCTTTAACTTGGGAAT 618
Qy      601  TGGCTTTGAAACTCAAGGCTAGAGGTGAGAGGAGGAGGAGGATTCATGATGAGAG 660
Db      619  TGGCTTTGAAACTCAAGGCTAGAGGTGAGAGGAGGAGGAGGATTCATGATGAGAG 678
Qy      661  TGAATGTGAGATATGAGAAACATCGATGCGGAGAGGAGGCTCTTGGGTTAACT 720
Db      679  TGAATGTGAGATATGAGAAACATCGATGCGGAGAGGAGGCTCTTGGGTTAACT 738
Qy      721  GACGCTCATGACGAAAGCGTGGGAGGCAAAAGGATTAGTACCCTGATGACGCGC 780
Db      739  GACGCTCATGACGAAAGCGTGGGAGGCAAAAGGATTAGTACCCTGATGACGCGC 798
Qy      781  CTAATACGATGCTACTAGTTGTGGGCTTATATGAGCTTTGGTAAGCAAGCTTAACGCGTGA 840
Db      799  CTAATACGATGCTACTAGTTGTGGGCTTATATGAGCTTTGGTAAGCTTAACGCGTGA 858
Qy      841  AGTTGACCGGCTGGGAGATACGATGCAAGATTAATACTCAAGGAATTGACGGGGAACCC 900
Db      859  AGTTGACCGGCTGGGAGATACGATGCAAGATTAATACTCAAGGAATTGACGGGGAACCC 918
```

```

Qy      901  GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGGGAAAAACCTTACTACCTT 960
Db      919  GCACAAGCGGTGATTAATGATGATTAATTCGATGCAACGGGAAAAACCTTACTACCTT 978
Qy      961  GACATGTAGCGAATTTTCTAGATAGATTAGTG-CTTCGGGAAACGCTTAACACAGTGCT 1019
Db      979  GACATGTAGCGAATTTTCTAGATAGATTAGTGCTTCGGGAAACGCTTAACACAGTGCT 1038
Qy      1020  GCATGCGTGTCTCACTCGTGTGTGAGATGTGGGTAAAGTCCCGCAACGACCGCAAC 1079
Db      1039  GCATGCGTGTCTCACTCGTGTGTGAGATGTGGGTAAAGTCCCGCAACGACCGCAAC 1098
Qy      1080  CCTGTGATTAATTGGCATCATTTGTTGGGCACTTTAATGAGACTCGCGGTGACAAAC 1139
Db      1099  CCTGTGATTAATTGGCATCATTTGTTGGGCACTTTAATGAGACTCGCGGTGACAAAC 1158
Qy      1140  GAGAGAAAGTGGGAGATGACGTAAGTCTCATGAGCTCTTAATGAGTGGGTAAAGGCTTCA 1199
Db      1159  GAGAGAAAGTGGGAGATGACGTAAGTCTCATGAGCTCTTAATGAGTGGGTAAAGGCTTCA 1218
Qy      1200  ATACATGCGCGCTGACAGAGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAAAGCGC 1259
Db      1219  ATACATGCGCGCTGACAGAGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAAAGCGC 1278
Qy      1260  GTCTGATCCGGAATCGGAGTCTGCACTCCGCTCGTAAGTGGGAATTCGTAGTAATCG 1319
Db      1279  GTCTGATCCGGAATCGGAGTCTGCACTCCGCTCGTAAGTGGGAATTCGTAGTAATCG 1338
Qy      1320  CGGATCAGCATGTGCGCGGTGAATAGTTCGCGGCTCTTGTATACACCGCCCGTACACCA 1379
Db      1339  CGGATCAGCATGTGCGCGGTGAATAGTTCGCGGCTCTTGTATACACCGCCCGTACACCA 1398
Qy      1380  TGGAGTGGGTTTACCAAGAAAGCAGTAGTCTAACGTAAGAGAGGCGCTTGGCAGGTG 1439
Db      1399  TGGAGTGGGTTTACCAAGAAAGCAGTAGTGTCTAACCTCGGAGAGAGCGTTAACCGAGT 1458
Qy      1440  AGATTCAATGACTGGGG 1455
Db      1459  TGTCAATGACTTGGG 1474

RESULT 13
US-10-659-948A-18
; Sequence 18, Application US/10659948A
; Publication No. US20040101946A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneq, Timothy A
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-294309
; CURRENT APPLICATION NUMBER: US/10/659,948A
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 18
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: N. Aestuarii-like AOB Paclione42 16S rDNA
US-10-659-948A-18

Query Match      93.3%; Score 1358.8; DB 7; Length 1467;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 37; Indels 3; Gaps 3;
```

QY 1 ATTGAACCTGGCGGCATGCTTTTACATGCAAGTCGAAACGCGACGAGTCTTGAT 60  
 DB 16 ATTGAACCTGGCGGCATGCTTTTACATGCAAGTCGAAACGCGACGAGTCTTGAC 75  
 QY 61 CTGTGGCGAGTGGCGGACGGGTGATATGCAATGGAAAGTATCTCAAGAAAGGGGGTA 120  
 DB 76 CTGTGGCGAGTGGCGGACGGGTGATATGCAATGGAAAGTATCTCAAGAAAGGGGGTA 135  
 QY 121 ACGCATCGAAAGATGTCTAATACCGCATATCTCTAAGAGAAAGAGGAGATCGAAA 180  
 DB 136 ACGCATCGAAAGATGTCTAATACCGCATATCTCTAAGAGAAAGAGGAGATCGAAA 195  
 QY 181 GACCTTGGCGCTTTTGAAGCGCGCCGATGTCTGATTAAGTATGTTGGTGAAGGCTTAC 240  
 DB 196 GACCTTGGCGCTTTTGAAGCGCGCCGATGTCTGATTAAGTATGTTGGTGAAGGCTTAC 255  
 QY 241 CAAGGCGAGATCAGTATGTTGCTTGAGAGACGACACGCTGAGACTGAGACACG 300  
 DB 256 CAAGGCGAGATCAGTATGTTGCTTGAGAGACGACACGCTGAGACTGAGACACG 315  
 QY 301 GCCCAGACTCTTACCGGAGGACGACGATGGGGAATTTTGGCAATGGGCGCAAGCTGATC 360  
 DB 316 GCCCAGACTCTTACCGGAGGACGACGATGGGGAATTTTGGCAATGGGCGCAAGCTGATC 375  
 QY 361 CAGCATGCGCGGTGAGTGAAGAGGCTTCGGGTTTGAAGCTCTTTTCAGTGAAGAGA 420  
 DB 376 CAGCATGCGCGGTGAGTGAAGAGGCTTCGGGTTTGAAGCTCTTTTCAGTGAAGAGA 434  
 QY 421 AAGGTTACGTTAAATATCTGATCATGACGGTATGCAAGAAAGACACCGGCTAAC 480  
 DB 435 AAGGTTACGTTAAATATCTGATCATGACGGTATGCAAGAAAGACACCGGCTAAC 494  
 QY 481 TACGTGCCAGACCGCGGTAATCTTAAGGTCGCAAGCGCTTAATCGGAATTCGGGCGT 540  
 DB 495 TACGTGCCAGACCGCGGTAATCTTAAGGTCGCAAGCGCTTAATCGGAATTCGGGCGT 554  
 QY 541 AAGGTTGCGCAGCGGCTTTTGAATCAGATGTGAATCCCGGCTTAACTTGGGAAT 600  
 DB 555 AAGGTTGCGCAGCGGCTTTTGAATCAGATGTGAATCCCGGCTTAACTTGGGAAT 614  
 QY 601 TGGGTTGAAACTACAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTACAG 660  
 DB 615 TGGGTTGAAACTACAGGCTAGAGTGTGGCAGAGGAGGTGGAATTCATGTGTACAG 674  
 QY 661 TGAATGCGTGAAGATATGGAAGAAATCATGATGCGGACACCTCTCGGGTTACACT 720  
 DB 675 TGAATGCGTGAAGATATGGAAGAAATCATGATGCGGACACCTCTCGGGTTACACT 734  
 QY 721 GACGCTCATGCAAGAAAGCGTGGGAGCAAAAGATTAATACCTGTGATCCAGCGC 780  
 DB 735 GACGCTCATGCAAGAAAGCGTGGGAGCAAAAGATTAATACCTGTGATCCAGCGC 794  
 QY 781 CTAAAGATGTCAACTAGTGTGTGGGCTTATTAGGCTTGTGAAGAAAGCTAACCGGTGA 840  
 DB 795 CTAAAGATGTCAACTAGTGTGTGGGCTTATTAGGCTTGTGAAGAAAGCTAACCGGTGA 854  
 QY 841 AGTTGACCGCTTGGGAGATACGGTCCGAGATTTAACTCAAGAAATGACGGGAGACC 900  
 DB 855 AGTTGACCGCTTGGGAGATACGGTCCGAGATTTAACTCAAGAAATGACGGGAGACC 914  
 QY 901 GCAAGAGCGGTGATTAATGAGATTAATGATGCAACGGGAAACCTTAACCTACCTT 960  
 DB 915 GCAAGAGCGGTGATTAATGAGATTAATGATGCAACGGGAAACCTTAACCTACCTT 974  
 QY 961 GACATGAGCGAATTTTCTAGAGATAGATTAGTGC-TTCGGGAAAGCTTAACAGAGTGT 1019  
 DB 975 GACATGAGCGAATTTTCTAGAGATAGATTAGTGC-TTCGGGAAAGCTTAACAGAGTGT 1034  
 QY 1020 GCATGCGTGTCTCAGCTCGTGTCTGATGATGTTGGGTTAACTCCCGCAACGAGCGAAC 1079  
 DB 1035 GCATGCGTGTCTCAGCTCGTGTCTGATGATGTTGGGTTAACTCCCGCAACGAGCGAAC 1094

QY 1080 CCTTGTCATTAAATGCCATCATTTGTGGGCACTTTAATGAGACTGCGGTGACAAACC 1139  
 DB 1095 CTTGTCAATTAATGTCATCATTTAGTTGGGCACTTTAATGAGACTGCGGTGACAAACC 1154  
 QY 1140 GGAAGAGTGGGGATGACGTCAGATCTTCATGAGCCCTTATGGTATAGGCTTTCACAGTA 1139  
 DB 1155 GGAAGAGTGGGGATGACGTCAGATCTTCATGAGCCCTTATGGTATAGGCTTTCACAGTA 1214  
 QY 1200 ATCAATAGGCGCGTACAGAGGTTGCCAACCCCGGAGGGGAGGTAACTCAGAAAGCGC 1259  
 DB 1215 ATCAATAGGCGCGTACAGAGGTTGCCAACCCCGGAGGGGAGGTAACTCAGAAAGCGC 1274  
 QY 1260 GTCCGATGCGGATCGGAGTCTGCAATCGACTCCGTAAGTCCGAATCGCTAGTAATCG 1319  
 DB 1275 GTCCGATGCGGATCGGAGTCTGCAATCGACTCCGTAAGTCCGAATCGCTAGTAATCG 1334  
 QY 1320 CGGATCAGCATGTGCGCGGTGAATAGTTCGGGCTTTGTACACACCGCCCGTCAACCA 1379  
 DB 1335 CGGATCAGCATGTGCGCGGTGAATAGTTCGGGCTTTGTACACACCGCCCGTCAACCA 1394  
 QY 1380 TGGAGTGGGTTTTCACAGAAAGCAGTATGTTAACTGTAAGGAGGCGCTTGGCAGGTC 1439  
 DB 1395 TGGAGTGGGTTTTCACAGAAAGCAGTATGTTAACTGTAAGGAGGCGCTTGGCAGGTC 1453  
 QY 1440 AGATTCAATGACTGG 1453  
 DB 1454 AGATTCAATGACTGG 1467

RESULT 14  
 US-10-659-980a-18  
 ; Sequence 18, Application US/10659980A  
 ; Publication No. US20040106133A1

; GENERAL INFORMATION:  
 ; APPLICANT: Hovaneq, Timothy A  
 ; TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria  
 ; FILE REFERENCE: 81289-284781  
 ; CURRENT FILING DATE: 2003-09-10  
 ; PRIOR APPLICATION NUMBER: US 09/573,684  
 ; PRIOR FILING DATE: 2000-05-19  
 ; PRIOR APPLICATION NUMBER: US 60/386,217  
 ; PRIOR FILING DATE: 2002-09-19  
 ; PRIOR APPLICATION NUMBER: US 60/386,218  
 ; PRIOR FILING DATE: 2002-09-19  
 ; PRIOR APPLICATION NUMBER: US 60/386,219  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: Patencin version 3.2  
 ; SEQ ID NO 18  
 ; LENGTH: 1467  
 ; TYPE: DNA  
 ; ORGANISM: Unknown  
 ; OTHER INFORMATION: N. Aestuarii-like AOB P4clone42 16S rDNA  
 ; US-10-659-980a-18

Query Match 93.3%; Score 1358.8; DB 7; Length 1467;  
 Best Local Similarity 97.2%; Pred. No. 0;  
 Matches 1414; Conservative 0; Mismatches 37; Indels 3; Gaps 3;

QY 1 ATTGAACGCTGGCGGCATGCTTTTACATGCAAGTCGAAACGCGACGAGTCTTGAT 60  
 DB 16 ATTGAACGCTGGCGGCATGCTTTTACATGCAAGTCGAAACGCGACGAGTCTTGAC 75  
 QY 61 CTGTGGCGAGTGGCGGACGGGTGATATGCAATGGAAAGTATCTCAAGAAAGGGGGTA 120  
 DB 76 CTGTGGCGAGTGGCGGACGGGTGATATGCAATGGAAAGTATCTCAAGAAAGGGGGTA 135  
 QY 121 ACGCATCGAAAGATGTCTAATACCGCATATCTCTAAGAGAAAGAGGAGATCGAAA 180  
 DB 136 ACGCATCGAAAGATGTCTAATACCGCATATCTCTAAGAGAAAGAGGAGATCGAAA 195

181 GACCTTGCGCTTTTGGAGCGCCGATGCTGATTAAGTAGTTGTTGGGTTAAAGCCCTAC 240  
186 GACCTTGCTGCTTTTGGAGCGCCGATGCTGATTAAGTAGTTGTTGGGTTAAAGCCCTAC 255  
241 CAAGCGCAGATCAGTAGTTGGTCTGAGAGACGACGACCACTGCGGACTGAGACACG 300  
246 CAAGCAACGATCAGTAGTTGGTCTGAGAGACGACGACCACTGCGGACTGAGACACG 315  
301 GCCCGACTCTTACCGGAGGACGAGTGGGAAATTTTGGACAAATGGGCGCAAGCCTGATC 360  
316 GCCCGACTCTTACCGGAGGACGAGTGGGAAATTTTGGACAAATGGGCGCAAGCCTGATC 375  
361 CAGCAATGCGCGGTGAGTGAAGAAGCCCTGCGGTTTGAAGCTCTTTCAGTCGAGAAGA 420  
376 CAGCAATGCGCGGTGAGTGAAGAAG - CTTGCGGTTTGAAGCTCTTTCAGTCGAGAAGA 434  
421 AAAGGTTACGTTAAATTAATCGTACTCATGACGGTATCGACGAAAGAACCGGCTAAC 480  
435 AAAGGTTGATTAATTAATCAACAATTATGATGTACGACAGAAAGAACCGGCTAAC 494  
481 TACGTCGACGACCGCCGGTATAGTAGGGTGCAGGCTTAATCGGAATTACTGGGGCT 540  
495 TACGTCGACGACCGCCGGTATAGTAGGGTGCAGGCTTAATCGGAATTACTGGGGCT 554  
541 AAAGGTCGCGACGCGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACTGGGAAT 600  
555 AAAGGTCGCGACGCGCTTTGTAAGTCAGATGTGAATCCCGGGCTTAACTGGGAAT 614  
601 TCGGTTTGAACCTACAAAGGCTAGAGTGTGCGACGAGGAGTGAATTCATGTGTAGCAG 660  
615 TCGGTTTGAACCTACAAAGGCTAGAGTGTGCGACGAGGAGTGAATTCATGTGTAGCAG 674  
661 TGAATGTGCTAGAGTATGGAAGAACTTCGATGTGGGAAGGCGCTCCCTGGTTAACT 720  
675 TGAATGTGCTAGAGTATGGAAGAACTTCGATGTGGGAAGGCGCTCCCTGGTTAACT 734  
721 GACGCTCATGACGAAAGCGTGGGAGCAAGGATTAGATACCCTGGTAGTCCAGCC 780  
735 GACGCTCATGACGAAAGCGTGGGAGCAAGGATTAGATACCCTGGTAGTCCAGCC 794  
781 CTAAACGATGTCATAGTTGTTGGGCTTATTAAGCTTGGTGAACGAACTAACCGGTGA 840  
795 CTAAACGATGTCATAGTTGTTGGGCTTATTAAGCTTGGTGAACGATGCTAACCGGTGA 854  
841 AGTTACCGGCTGGGAGTACGCTGCGCAAGATTAATACTCAAGAAATTTGACGGGAGACC 900  
855 AGTTACCGGCTGGGAGTACGCTGCGCAAGATTAATACTCAAGAAATTTGACGGGAGACC 914  
901 GCACAAAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTCT 960  
915 GCACAAAGCGGTGATTAATGATTAATTCGATGCAAGCGGAAACCTTACCTACCTCT 974  
961 GACATGTAGCGAATTTTCTAGAGATAGATTAGTC - TTCCGGAACGCTTAACACAGGTCT 1019  
975 GACATGTAGCGAATTTTCTAGAGATAGATTAGTC - TTCCGGAACGCTTAACACAGGTCT 1034  
1020 GCATGGCTGTGTCAGCTGCTGTGCTGAGATGTTGGGTTAACTCCCGCAAGAGCGCAAC 1079  
1035 GCATGGCTGTGTCAGCTGCTGTGCTGAGATGTTGGGTTAACTCCCGCAAGAGCGCAAC 1094  
1080 CTTTGTCAATTAATGCGCATCATTTGTTGGGCACTTTATGAGACTGCGCGGTGACAAAC 1139  
1095 CTTTGTCAATTAATGCGCATCATTTGTTGGGCACTTTATGAGACTGCGCGGTGACAAAC 1154  
1140 GGAGGAAGGTGGGATGACGTCAGTCCATGAGCCCTTATGGGTAGGGCTTACACAGTA 1199  
1155 GGAGGAAGGTGGGATGACGTCAGTCCATGAGCCCTTATGGGTAGGGCTTACACAGTA 1214  
1200 ATACATAGCGCGCTACAGAGGTTGCGCAACCCGCGAGGGGAGCTTAATCTCAAGAAAGCG 1259  
1215 ATACATAGCGCGCTACAGAGGTTGCGCAACCCGCGAGGGGAGCTTAATCTCAAGAAAGCG 1274  
1260 GTCTGATGTCGGATGGAGTCTGCAACTCGGTGAAGTGGGAATCGGATGTAATCG 1319

1275 GTCTGATGTCGGATCGGAGTGTGCAACTGCACTCGGTAAGTCGGAATCGCTAGTAATCG 1334  
1320 CGGATCAGCATGTCGGGCTGAATACGTTCCCGGCTCTTGTACACACCGCCCGTACACCA 1379  
1335 CGGATCAGCATGTCGGGCTGAATACGTTCCCGGCTCTTGTACACACCGCCCGTACACCA 1394  
1380 TGGAGTGGGTTTACCGCAAGACAGTAGTCTAACCGTAAAGAGGAGCGCTTGCCAGGTG 1439  
1395 TGGAGTGGGTTTACCGCAAGACAGTAGTCTAACCGTAAAGAGGAGCGCTTGCCAGGTG 1453  
1440 AGATTCACTGACTGG 1453  
1454 AGATTCACTGACTGG 1467  
RESULT 15  
US-10-659-983A-18  
; Sequence 18, Application US/10659983A  
; Publication No. US2004015731A1  
; GENERAL INFORMATION:  
; APPLICANT: Hovane, Timothy A  
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria  
; FILE REFERENCE: 81289-28479  
; CURRENT APPLICATION NUMBER: US/10/659, 983A  
; CURRENT FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: US 09/573, 684  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 60/386, 217  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386, 218  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: US 60/386, 219  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 1467  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: N. Aestuarii-like AOB PaClone42 16S rDNA  
US-10-659-983A-18  
Query Match 93.3%; Score 1358.8; DB 7; Length 1467;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1414; Conservative 0; Mismatches 37; Indels 3; Gaps 3;  
1 ATTGAACGCTGGCGCATGCTTTACACATGCAAGTCGAACCGGACGACGATGCTTGAT 60  
16 ATTGAACGCTGGCGCATGCTTTACACATGCAAGTCGAACCGGACGACGATGCTTGAT 75  
61 CTGTGTGCGAAGTGGCGGACGCGGTGATGATGCAATCGGAACGTAATCCAGAAAGGGGGTA 120  
76 CTGTGTGCGAAGTGGCGGACGCGGTGATGATGCAATCGGAACGTAATCCAGAAAGGGGGTA 135  
121 ACGCATGGAAGAGTGTCTAATACCGCATTAATCTTAAGAGAGAAACAGAGGGGATCGAA 180  
136 ACGCATGGAAGAGTGTCTAATACCGCATTAATCTTAAGAGAGAAACAGAGGGGATCGAA 195  
181 GACCTTGCGCTTTTGGAGCGCCGATGCTGATTAAGTGAATGTTGGTGGGTTAAAGCCCTAC 240  
196 GACCTTGCGCTTTTGGAGCGCCGATGCTGATTAAGTGAATGTTGGTGGGTTAAAGCCCTAC 255  
241 CAAGCGCAGATCAGTAGTTGGTCTGAGAGACGACGACCACTGCGGACTGAGACACG 300  
256 CAAGCGCAGATCAGTAGTTGGTCTGAGAGACGACGACCACTGCGGACTGAGACACG 315  
301 GCCCGACTCTTACCGGAGGACGAGTGGGAAATTTTGGACAAATGGGCGCAAGCCTGATC 360  
316 GCCCGACTCTTACCGGAGGACGAGTGGGAAATTTTGGACAAATGGGCGCAAGCCTGATC 375  
361 CAGCAATGCGCGGTGAGTGAAGAAGCCTTCCGGTTTGAAGCTCTTTCAGTCGAGAAGA 420

DB 376 CAGCAATCCCGCTAGTGAGAGAG-CTTCCGGTGTAAAGCTTTTCAGTCGAGAGAA 434  
QY 421 AAAAGTTACGGTAATTAATCGTACCTGACGATGACGATGACAGAGAAACACCGGCTAAC 480  
DB 435 AAAAGTTGACTTAATTAATCAACACTTAATGATGATGCCAGAGAGAGACCGGCTAAC 494  
QY 481 TACGTCCAGCAGCCCGGCTAATACGTAGGGTGCAGCGCTTAATCGGAATTACTGGCGCT 540  
DB 495 TACGTCCAGCAGCCCGGCTAATACGTAGGGTGCAGCGCTTAATCGGAATTACTGGCGCT 554  
QY 541 AAAAGTGCAGAGCGCGCTTTGTAAGTCAAGATGTGAATCCCGGGCTTAACCTGGGAAT 600  
DB 555 AAAAGTGCAGAGCGCGCTTTGTAAGTCAAGATGTGAATCCCGGGCTTAACCTGGGAAT 614  
QY 601 TCCGTTTGAATTAACAAGCTAGTGTGCGAGAGGAGTGAATTCATGATGTAGCAG 660  
DB 615 TCCGTTTGAATTAACAAGCTAGTGTGCGAGAGGAGTGAATTCATGATGTAGCAG 674  
QY 661 TGAATTCGCTAGAGATATGGAAGAACATGATGCGAGAGCGCTCTGGGTTAACACT 720  
DB 675 TGAATTCGCTAGAGATATGGAAGAACATGATGCGAGAGCGCTCTGGGTTAACACT 734  
QY 721 GACGCTGATGACAGAAAGCTGGGGAGCAACAGATTAATGATACCTGGTATCCAGCGC 780  
DB 735 GACGCTGATGACAGAAAGCTGGGGAGCAACAGATTAATGATACCTGGTATCCAGCGC 794  
QY 781 CTAAACGATGCTAACTAGTTGTGGGCTTAATTAAGCTTGTATGACGATTAACCGGCTGA 840  
DB 795 CTAAACGATGCTAACTAGTTGTGGGCTTACTAGGCTTGTATGACGATTAACCGGCTGA 854  
QY 841 AGTTGACCGCTGGGGAGTACGGTCCGCAAGATTAATACTCAAGAGATTGACGGGAGCC 900  
DB 855 AGTTGACCGCTGGGGAGTACGGTCCGCAAGATTAATACTCAAGAGATTGACGGGAGCC 914  
QY 901 GCACAAGCGGTGATTAATGATTAATGATGATGCAACGCGAAAACCTTACCTTACCCTT 960  
DB 915 GCACAAGCGGTGATTAATGATTAATGATGATGCAACGCGAAAACCTTACCTTACCCTT 974  
QY 961 GACATGTACGGAATTTTCTAGAGATAGATTAATGATGATGCAACGCGAAAACCTTACCTT 1019  
DB 975 GACATGTACGGAATTTTCTAGAGATAGATTAATGATGATGCAACGCGAAAACCTTACCTT 1034  
QY 1020 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079  
DB 1035 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1094  
QY 1080 CCTTGTCAATTAATTTGCTATTTGCTTGGGCACTTAAATGAGATGCTGCTGCTGCTGCTGCT 1139  
DB 1095 CCTTGTCAATTAATTTGCTATTTAGTTGGGCACTTAAATGAGATGCTGCTGCTGCTGCTGCT 1154  
QY 1140 GAGAGAAAGTGGGAGTGAAGTCAAGTCTTATGAGCCCTTAATGGGTAGGGCTTCAACGTA 1199  
DB 1155 GAGAGAAAGTGGGAGTGAAGTCAAGTCTTATGAGCCCTTAATGGGTAGGGCTTCAACGTA 1214  
QY 1200 ATACAAATGGGCGGTGAGAGAGGTTGGCAACCGGAGAGGGGAGTAACTCAGAAAAGCGC 1259  
DB 1215 ATACAAATGGGCGGTGAGAGAGGTTGGCAACCGGAGAGGGGAGTAACTCAGAAAAGCGC 1274  
QY 1260 GTGCTAGTCCGGAATCGGAGTCTGCAACTGCACTCCGTAAGTGGGAATGCTAGTAATCG 1319  
DB 1275 GTGCTAGTCCGGAATCGGAGTCTGCAACTGCACTCCGTAAGTGGGAATGCTAGTAATCG 1334  
QY 1320 CGGATCAGCATGTCGCGGTGAATTCGTTCCCGGCTTGTGACACACCGCCCGTACACCA 1379  
DB 1335 CGGATCAGCATGTCGCGGTGAATTCGTTCCCGGCTTGTGACACACCGCCCGTACACCA 1394  
QY 1380 TGGGAGTGGGTTTCAACAGAGAGAGTATCTAACCGTAAGAGAGGCGCTTGGCCACGCTG 1439  
DB 1395 TGGGAGTGGGTTTCAACAGAGAGAGTATCTAACCGTAAGAGAGGCGCTTGGCCACGCTG 1453  
QY 1440 AGATTCAATGACTGG 1453

DB 1454 AGATTCAATGACTGG 1467  
RESULT 16  
US-10-659-948A-3  
Sequence 3, Application US/10659948A  
Publication No. US20040101946A1  
GENERAL INFORMATION:  
APPLICANT: Hovaneq, Timothy A  
TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria  
FILE REFERENCE: 81289-294309  
CURRENT APPLICATION NUMBER: US/10/659,948A  
CURRENT FILING DATE: 2003-09-10  
PRIOR APPLICATION NUMBER: US 09/573,684  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 60/386,217  
PRIOR FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: US 60/386,218  
PRIOR FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: US 60/386,219  
PRIOR FILING DATE: 2002-09-19  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 3  
LENGTH: 1458  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: AOB Type B R3clones 168 rDNA  
US-10-659-948A-3  
Query Match 88.4%; Score 1287.6; DB 7; Length 1458;  
Best Local Similarity 93.6%; Pred. No. 0;  
Matches 1366; Conservative 0; Mismatches 89; Indels 5; Gaps 2;  
QY 1 ATTTGAACGCTGGCGGATAGCTTTTACATGATGCAAGTGCAGAACGCGACAGCGATGCTTGCAT 60  
DB 1 ATTTGAACGCTGGCGGATAGCTTTTACATGATGCAAGTGCAGAACGCGACAGCGGCGCAAC-- 58  
QY 61 CTGTGTGCGAGTGGCGGAGCGGCTGATATGATCGGAACGTAATCAGAAAGGGGGGTA 120  
DB 59 CTGTGTGCGAGTGGCGGAGCGGCTGATATGATCGGAACGTAATCAGAAAGGGGGGTA 118  
QY 121 ACGCATGCAAGATGTCTAATACCGCATATACTTAAGAGAGAAAGCGGGGATGCGAA 180  
DB 119 ACGCATGCAAGATGTGTCTAATACCGCATATCTCCACGAGAGAAAGCGGGGATGCGAA 178  
QY 181 GACCTTGCCTTTTGAAGCGCGCGATGCTGATTTAGCTAGTTGGTGGGTTAAAGGCTTAC 240  
DB 179 GACCTTGCCTTTTGAAGCGCGCGATGCTGATTTAGCTAGTTGGTGGGTTAAAGGCTTAC 238  
QY 241 CAAGCGAGATCAGTAATGCTGAGAGAGAGCAACCACTGGAATGAGACAGC 300  
DB 239 CAAGCGAGATCAGTAATGCTGAGAGAGAGCAACCACTGGAATGAGACAGC 298  
QY 301 GCCCAGACTCTTAACGGGAGGAGAGTGGGAAATTTTGAACAATGGGCGCAAGCTTGATC 360  
DB 299 GCCCAGACTCTTAACGGGAGGAGAGTGGGAAATTTTGAACAATGGGCGCAAGCTTGATC 358  
QY 361 CAGCAATGCCGCTGATGATGAAGAGGCTTCCGGTTGTAAGCTCTTTCAGTGAAGA 420  
DB 359 CAGCAATGCCGCTGATGATGAAGAGGCTTCCGGTTGTAAGCTCTTTCAGTGAAGA 418  
QY 421 AAAAGTTAGGTAATTAATCGTACCTCATGACGTAATGCAAGAGAGCAACCGGCTAAC 480  
DB 419 AAAAGTTAGGTAATTAATCGTACCTCATGACGTAATGCAAGAGAGCAACCGGCTAAC 478  
QY 481 TACGTCCAGCAGCCCGGCTAATACGTAGGGTGCAGCGCTTAATCGGAATTACTGGCGCT 540  
DB 479 TACGTCCAGCAGCCCGGCTAATACGTAGGGTGCAGCGCTTAATCGGAATTACTGGCGCT 538  
QY 541 AAAAGTGCAGAGCGCGCTTTGTAAGTCAAGATGTGAATCCCGGGCTTAACCTGGGAAT 600



```

OY 781 CTAACGATGTCACATAGTGTGGCCCTTATTAGCCTTGTAACGACCTAACCGCTGA 840
DB 779 CTAACGATGTCACATAGTGTGGCCCTTATTAGCCTTGTAACGACCTAACCGCTGA 838
OY 841 AGTTGACCGCTGGGGAGTACCGTCCGCAAGATTAAACCTCAAGAAATGACGGGGACCC 900
DB 839 AGTTGACCGCTGGGGAGTACCGTCCGCAAGATTAAACCTCAAGAAATGACGGGGACCC 898
OY 901 GCAACAGCGGTGATTAATGATGATTAATTCATGACACCGCAAAACCTTACCTACCTT 960
DB 899 GCAACAGCGGTGATTAATGATGATTAATTCATGACACCGCAAAACCTTACCTACCTT 958
OY 961 GACATGTACGCAATTTTCTAGAGATGATTAATGATGATGATGATGATGATGATGATG 1017
DB 959 GACATGTACGCAATTTTCTAGAGATGATTAATGATGATGATGATGATGATGATGATG 1018
OY 1018 CTGCAATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077
DB 1019 CTGCAATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1078
OY 1078 ACCCTGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1137
DB 1079 ACCCTGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1138
OY 1138 CCGGAGGAGAGTGGGGAGTACGTCATGATGATGATGATGATGATGATGATGATGATG 1197
DB 1139 CCGGAGGAGAGTGGGGAGTACGTCATGATGATGATGATGATGATGATGATGATGATG 1198
OY 1198 TAATACATAGCGCGTACAGAGGCTTCCACCCCGAGGGGAGCTTAATCTCAAGAAC 1257
DB 1199 TAATACATAGCGCGTACAGAGGCTTCCACCCCGAGGGGAGCTTAATCTCAAGAAC 1258
OY 1258 GCGTGTATGCTGGGATGCGAGTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317
DB 1259 GCGTGTATGCTGGGATGCGAGTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1318
OY 1318 CCGGATGACATGCTGCGGCTGATACGTTCCCGGCTTGTACACACCGCCGCTGACAC 1377
DB 1319 CCGGATGACATGCTGCGGCTGATACGTTCCCGGCTTGTACACACCGCCGCTGACAC 1378
OY 1378 CATGGAGATGGGTTTCCACAGAGAGTGTCTAACCGTAAAGAGGGCCCTTCCACCG 1437
DB 1379 CATGGAGATGGGTTTCCACAGAGAGTGTCTAACCGTAAAGAGGGCCCTTCCACCG 1438
OY 1438 TGAGATTCACTGAGGCTG 1457
DB 1439 TGAGATTCACTGAGGCTG 1458

RESULT 18
US-10-659-983A-3
; Sequence 3, Application US/10659983A
; Publication No. US2004015731A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneq, Timochy A
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284779
; CURRENT APPLICATION NUMBER: US/10/659, 983A
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1458
; TYPE: DNA
```

```

; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type B R3clones 16S rDNA
US-10-659-983A-3

Query Match      88.4%; Score 1287.6; DB 7; Length 1458;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1366; Conservative 0; Mismatches 89; Indels 5; Gaps 2;

OY 1 ATTGAACCTGCGCGCATGCTTTTACACATGCAAGTCCGAACCGGACGACGAGTGTGAT 60
DB 1 ATTGAACCTGCGCGCATGCTTTTACACATGCAAGTCCGAACCGGACGACGAGTGTGAT 58
OY 61 CTGGTGGGAGTGGGGAGCGGGTGTATGATGATGATGATGATGATGATGATGATGATG 120
DB 59 CTGGTGGGAGTGGGGAGCGGGTGTATGATGATGATGATGATGATGATGATGATGATG 118
OY 121 ACGCATCGAAAGATGTGCTAATACCGCATATCTCTAAGAGGAAAGAGGGGATCGAAA 180
DB 119 ACGCATCGAAAGATGTGCTAATACCGCATATCTCTAAGAGGAAAGAGGGGATCGCAA 178
OY 181 GACCTTGGCTTTTGGAGCGCGCATGCTGATTAATGCTAATGCTGCTGCTGCTGCTG 240
DB 179 GACCTTGGCTTTTGGAGCGCGCATGCTGATTAATGCTAATGCTGCTGCTGCTGCTG 238
OY 241 CAAGCGACGATCACTAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 239 CAAGCGACGATCACTAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 298
OY 301 GCCCAAGCTCTTACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 299 GCCCAAGCTCTTACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358
OY 361 CAGCAATGCGCGGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 359 CAGCAATGCGCGGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418
OY 421 AAGGTTAAGGTAATTAATCGTACATGACGATGATGATGATGATGATGATGATGATG 480
DB 419 AAGGTTAAGGTAATTAATCGTACATGACGATGATGATGATGATGATGATGATGATG 478
OY 481 TAGGTGCGAGAGCGCGGTAAATACGTAAGGAGTGAAGCGTTAATCGTAATCTGAGCG 540
DB 479 TAGGTGCGAGAGCGCGGTAAATACGTAAGGAGTGAAGCGTTAATCGTAATCTGAGCG 538
OY 541 AAGGTTGCGAGAGCGCGGTAAATACGTAAGGAGTGAAGCGTTAATCGTAATCTGAGCG 600
DB 539 AAGGTTGCGAGAGCGCGGTAAATACGTAAGGAGTGAAGCGTTAATCGTAATCTGAGCG 598
OY 601 TGGCTTTGAAATCTAAGGCTGAGTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 599 TGGCTTTGAAATCTAAGGCTGAGTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
OY 661 TGAATATGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 659 TGAATATGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
OY 721 GACGCTCATGCAAGAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 719 GACGCTCATGCAAGAGCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778
OY 781 CTAACGATGTCACATAGTGTGGCCCTTATTAGCCTTGTAACGACCTAACCGCTGA 840
DB 779 CTAACGATGTCACATAGTGTGGCCCTTATTAGCCTTGTAACGACCTAACCGCTGA 838
OY 841 AGTTGACCGCTGGGGAGTACCGTCCGCAAGATTAAACCTCAAGAAATGACGGGGACCC 900
DB 839 AGTTGACCGCTGGGGAGTACCGTCCGCAAGATTAAACCTCAAGAAATGACGGGGACCC 898
OY 901 GCAACAGCGGTGATTAATGATGATTAATTCATGACACCGCAAAACCTTACCTACCTT 960
DB 899 GCAACAGCGGTGATTAATGATGATTAATTCATGACACCGCAAAACCTTACCTACCTT 958
```

QY	961	GACATGTACCGAATTTTCTTAGAGATAGATTAGTGC--TTGGGGAACGCTTAACACAGGTG	1017
Db	959	GACATGTATCCGAAAGCCCGCCCGAGAGGTGGGTGTGCCGAAAGGAGCGGTATACACAGGTG	1018
QY	1018	CTGCATGGCTGTGCGTCACTCGTGTGCGTAGATGTGTGGATTAAAGTCCGCGAACGAGCGCA	1077
Db	1019	CTGCATGGCTGTGCGTCACTCGTGTGCGTAGATGTGTGGATTAAAGTCCGCGAACGAGCGCA	1078
QY	1078	ACCCCTTGTCATTAATTATGGCATATTTGGTTGGGCATTAAATGACACTGCCGGTGCACAA	1137
Db	1079	ACCCCTTGTCATTAATTATGGCATATTTCAATTCAATTGGGCATTTAATGAACCTGCCGGTGCACAA	1138
QY	1138	CCGAGAGAGGTGGGAGTGAAGTCACAGTCTTCATATGSCCTTAATGSGTATGSCCTTCAACG	1197
Db	1139	CCGAGAGAGGTGGGAGTGAAGTCACAGTCTTCATATGSCCTTAATGSGTATGSGCTTCAACG	1198
QY	1198	TAAATCAATGGCGGCTACAGAGGTTGGCCAAACCGCGAGGGGAGCTAATCTCAGAAAGC	1257
Db	1199	TAAATCAATGGCGGCTTACAGAGGTTGGCCAAACCGCGAGGGGAGCTAATCTCAGAAAGC	1258
QY	1258	GCGTGTATGTCGGATTCGGAATCTTGCAACTGCACTCCGTTGAAGTGGAAATCGCATAGTAAT	1317
Db	1259	GCGTGTATGTCGGATTCGGAATCTTGCAACTGCACTCCGTTGAAGTGGAAATCGCATAGTAAT	1318
QY	1318	CGCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGCTTTGTATACACACCGCCGCTCACAC	1377
Db	1319	CGCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGCTTTGTATACACACCGCCCGCTCACAC	1378
QY	1378	CATGGAGTGGGTTTCAACAAGACAGTAAGTTAAACGTTAAGAGAGGGCGCTTGCACAGG	1437
Db	1379	CATGGAGTGGGTTTCAACAAGACAGTAAGTTAAACGTTAAGAGAGGGCGCTTGCACAGG	1438
QY	1438	TGAATTTCAATGATGGGGTGTG	1457
Db	1439	TGAATTTCAATGATGGGGTGTG	1458

```

RESULT 19
US-10-659-948A-4
; Sequence 4, Application US/10659948A
; Publication No. US20040101946A1
; GENERAL INFORMATION:
; APPLICANT: Hovaneq, Timothy A
; TITLE OF INVENTION: Method of Using Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-294309
; CURRENT APPLICATION NUMBER: US/10/659, 948A
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: MOB Type C R5clone47 16S rDNA
US-10-659-948A-4

```

Query Match	84.4%	Score 1230.2	DB 7	Length 1460
Best Local Similarity	92.3%	Pred. No. 0		
Matches 131	0	Mismatches 103	Indels 9	Gaps 5

Qy 1 ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGGCAGACCGCATGCTTGCA 60  
Db 1 ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGGCAG--CGGGGCTTCGCG 58

[illegible]



```
Db      1138  ||||| 1197
      1138  AAACCGGAGGAGTGGGGATGACGTCAAGTCTCATGGCCCTTATGGCTAGGGCTTCC 1197
      1138  ||||| 1254
      1138  AGCTAATACAAATGCGCCGACGAGGGTTCACACCCGAGGGAGGCTAATCTCAGA 1254
      1138  ||||| 1257
      1138  AGCTAATACAAATGCGCCGACGAGGGTTCACACCCGAGGGAGGCTAATCTCAGA 1257
      1138  ||||| 1314
      1138  AGCGGTCTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCC 1314
      1138  ||||| 1317
      1138  AGCAGCTCTAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCC 1317
      1138  ||||| 1374
      1138  AATCGCGATCAGCATGTCGCGGTGAATACGTTCCCGGCTTTTACACACCGCCCTCA 1374
      1138  ||||| 1377
      1138  AATCGCGATCAGCATGTCGCGGTGAATACGTTCCCGGCTTTTACACACCGCCCTCA 1377
      1138  ||||| 1434
      1138  CACCATGGAGTGGCTTTTACACGAGAGCATGATTAACCGTAAGAGAGGCGCTTGCA 1434
      1138  ||||| 1437
      1138  CACCATGGAGTGGCTTTTACACGAGAGCATGATTAACCGTAAGAGAGGCGCTTGCA 1437
      1138  ||||| 1457
      1138  CGGTGATTCATGACTGGGGT 1457
      1138  ||||| 1460
      1138  CGGTGATTCATGACTGGGGT 1460
      1138  ||||| 1460
      1138  CGGTGATTCATGACTGGGGT 1460

RESULT 20
US-10-659-980A-4
Sequence 4, Application US/10659980A
Publication No. US20040106133A1
GENERAL INFORMATION:
APPLICANT: Hovaneq, Timothy A
TITLE OF INVENTION: Method for Detecting Ammonia-Oxidizing Bacteria
FILE REFERENCE: 81289-284781
CURRENT APPLICATION NUMBER: US/10/659, 980A
CURRENT FILING DATE: 2003-09-10
PRIOR APPLICATION NUMBER: US 09/573,684
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 60/386,217
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: US 60/386,218
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: US 60/386,219
PRIOR FILING DATE: 2002-09-19
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4
LENGTH: 1460
TYPE: DNA
ORGANISM: Unknown
OTHER INFORMATION: AOB Type C R5clone47 16S rDNA
US-10-659-980A-4

Query Match      84.4%; Score 1230.2; DB 7; Length 1460;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1351; Conservative 0; Mismatches 103; Indels 9; Gaps 5;
```

```
Db      228  ||||| 297
      228  CAAGCAACGATCACTAGTGTCTGAGAGGACGACCAACCACTGGGACTGAGACAG 297
      228  ||||| 360
      228  GCCCAAGTCTCTACGAGAGGACGAGTGGGAATTTTGAACATGAGCGCAAGCTGATC 360
      228  ||||| 357
      228  GCCCAAGTCTCTACGAGAGGACGAGTGGGAATTTTGAACATGAGCGCAAGCTGATC 357
      228  ||||| 420
      228  CAGCAATGCGCGTGAAGTGAAGAGCCCTGGGGTGAAGCTCTTTCAGTCGAGAA 420
      228  ||||| 417
      228  CAGCAATGCGCGTGAAGTGAAGAGCCCTGGGGTGAAGCTCTTTCAGTCGAGAA 417
      228  ||||| 480
      228  AAGGTTACGTTAATTCGTAATCTGATCATGACGATTCGACAGAGAAAGCAACG 480
      228  ||||| 477
      228  AAGGTTACGTTAATTCGTAATCTGATCATGACGATTCGACAGAGAAAGCAACG 477
      228  ||||| 540
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 540
      228  ||||| 537
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 537
      228  ||||| 597
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 597
      228  ||||| 600
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 600
      228  ||||| 597
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 597
      228  ||||| 660
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 660
      228  ||||| 657
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 657
      228  ||||| 720
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 720
      228  ||||| 717
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 717
      228  ||||| 780
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 780
      228  ||||| 777
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 777
      228  ||||| 838
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 838
      228  ||||| 837
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 837
      228  ||||| 897
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 897
      228  ||||| 957
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 957
      228  ||||| 957
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 957
      228  ||||| 1015
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1015
      228  ||||| 1017
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1017
      228  ||||| 1075
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1075
      228  ||||| 1134
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1134
      228  ||||| 1137
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1137
      228  ||||| 1194
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1194
      228  ||||| 1197
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1197
      228  ||||| 1254
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1254
      228  ||||| 1257
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1257
      228  ||||| 1314
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1314
      228  ||||| 1317
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1317
      228  ||||| 1374
      228  TACGTGCGACGACCGCGGTAAATACGTAGGGTGCAGCGTTAATCGAATTAAGTGG 1374
```

```
Db      1318 AATCGGAGTCAAGCATGCCGCGGTGAATACGTTCCCGGCTTTGTAACACACCGCCGCTCA 1377
Qy      1375 CACCAATGGAGTGGGTTTCAACAGAAAGCATAGTCTAAACGTTAGAGAGGCGCTTGCCA 1434
Db      1378 CACCAATGGAGTGGGTTTCAACAGAAAGCATAGTCTAAACGTTAGAGAGAGCGCTTGCCA 1437
Qy      1435 CGGTGAGATTCAATGACTGGGGTG 1457
Db      1438 CGGTGGGGGTCAATGACTGGGGTG 1460

RESULT 21
US-10-659-983A-4
; Sequence 4, Application US/10659983A
; Publication No. US20040157313A1
; GENERAL INFORMATION:
; APPLICANT: Hovanec, Timothy A
; TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
; FILE REFERENCE: 81289-284779
; CURRENT APPLICATION NUMBER: US/10/659,983A
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: US 09/573,684
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 60/386,217
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,218
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/386,219
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AOB Type C Rclone47 16D rDNA
US-10-659-983A-4

Query Match      84.4%; Score 1230.2; DB 7; Length 1460;
Beet Local Similarity 92.3%; Pred. No. 0;
Matches 1351; Conservative 0; Mismatches 103; Indels 9; Gaps 5;

Qy      1 AATTGAACCTGGCGGATGCTTTACATGCAAGTCGAAACGGACACAGAGATGCTGCAT 60
Db      1 AATTGAACCTGGCGGATGCTTTACATGCAAGTCGAAACGGACAG--CGGGGGCTTCGAC 58
Qy      61 CTGGTGGAGTGGGCGGACGGGTGAGTATGATCGAAACGTATCCAGAAAGGGGGTA 120
Db      59 CTGGCGGGAGTGGGGAACGGGTGAGTATGATCGAAACGTATCCTTAAGTGGGGAATTA 118
Qy      121 ACGCATGCAAGATGTGCTAATACCGCATATATCTTAAAGAGAGAAAGCAGGGATCGAAA 180
Db      119 ACGCATGCAAGATGTGCTAATACCGCATAT-CTCTGAGAGAGAAAGCAGGGGATCGCAA 177
Qy      181 GACCTTTGGCTTTTGGACCGCGCGATGTCTGATTGATGCTAGTTGGTGGGTTAAAGCCCTAC 240
Db      178 GACCTTTGGCTTTAAAGACCGCGCGATGTCTGATTGATGCTAGTTGGTGGGTTAAAGCCCTAC 237
Qy      241 CAAGCGAGCATAGTATGCTGAGAGAGACGACCACTATGGGACTGAGACAG 300
Db      238 CAAGCGAGCATAGTATGCTGAGAGAGACGACCACTATGGGACTGAGACAG 297
Qy      301 GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATC 360
Db      298 GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCTGATC 357
Qy      361 CAGCAATGCCCGGTGAGTGAAGAGAGCCTTGGGGTTGTAAGCTTTTCAAGTCGAGAGA 420
Db      358 CAGCAATGCCCGGTGAGTGAAGAGAGCCTTGGGGTTGTAAGCTTTTCAAGAGA 417
Qy      421 AAGGTACGGTAAATATATCGTCACTGACGATGACGAGAGAGACCGGCTAAC 480
```

```
Db      418 AAGATCATGATGATATATATATGATTTATGACGGTACTGACAGAAAAGCACCAGCTAAC 477
Qy      481 TACGTGCCAGACGCCCGGTATATACGTATGGGTGCAAGGCTTAATTCGGAATTAATCTGGGCT 540
Db      478 TACGTGCCAGACGCCCGGTATATACGTATGGGTGCAAGGCTTAATTCGGAATTAATCTGGGCT 537
Qy      541 AAGGGTGGCAGAGCGCTTTGTAAAGTGTGAAATGCCCGGGCTTAACCTGGGAAT 600
Db      538 AAGGGTGGCAGAGCGCTTTGTAAAGTGTGAAATGCCCGGGCTTAACCTGGGAAT 597
Qy      601 TCGCTTTGAAACTACAAAGCTAGAGTGGGAGAGAGAGTGGAAATTCATGTGTAGCAG 660
Db      598 TCGCTTTGAAACTACAAAGCTAGAGTGGGAGAGAGAGTGGAAATTCATGTGTAGCAG 657
Qy      661 TGAATTCGTGAGATATGAAAGAACATCATGTCGGAAGGAGGCTCTCTGGGTAAACT 720
Db      658 TGAATTCGTGAGATATGAAAGAACATCATGTCGGAAGGAGGCTCTCTGGGTAAACT 717
Qy      721 GACGCTCATGACAGAAAGCGTGGGGAGCAACAGGATTAATACCTGTGTATGACGAC 780
Db      718 GACGCTCATGACAGAAAGCGTGGGGAGCAACAGGATTAATACCTGTGTATGACGAC 777
Qy      781 CTAAACGATGTCACTAGTTGTGGCTTATTA--GGCTTGTAGAGAGCTAACGCGT 838
Db      778 CTAAACGATGTCACTAGTTGTGGCTTATTAAGATTTTGTAGAGCTAACGCGT 837
Qy      839 GAAGTTGACCGCTGGGGAGTACGCTGCGAAGATTAATACTCAAGAGATTAACGCGGAC 898
Db      838 GAAGTTGACCGCTGGGGAGTACGCTGCGAAGATTAATACTCAAGAGATTAACGCGGAC 897
Qy      899 CCGCAGAGCGGTGATTAATGATTAATCGATGACGAGAAACCTTACCTACCC 958
Db      898 CCGCAGAGCGGTGATTAATGATTAATCGATGACGAGAAACCTTACCTACCC 957
Qy      959 TTGACATGTAGCGAATTTTCTAGAGATGATTAATGAC--TTCCGAGACGCTAACAGAG 1015
Db      958 TTGACATGTAGCGAATTTTCTAGAGATGATTAATGAC--TTCCGAGACGCTAACAGAG 1017
Qy      1016 TGCATGATGCTGTCTGTCAGCTGCTGTCTGTAGATGTTGGTTAACTCCGCAACGAGCG 1075
Db      1018 TGCATGATGCTGTCTGTCAGCTGCTGTCTGTAGATGTTGGTTAACTCCGCAACGAGCG 1077
Qy      1076 CAACCTTTGCTAATTAATTTGCATCATTT--TGGTGGGCACTTAATGAGACTGCGGATAC 1134
Db      1078 CAACCTTTGCTAATTAATTTGCATCATTTTAATGAGACTTAATGAGACTGCGGATAC 1137
Qy      1135 AAACCGAGAGAGTGGGATGACGTCAAGTCTCATGCGCTTATGAGTGGCTTAC 1194
Db      1138 AAACCGAGAGAGTGGGATGACGTCAAGTCTCATGCGCTTATGAGTGGCTTAC 1197
Qy      1195 ACGTAATACATGGCGGTACAGAGGTTGCGCAACCCCGAGGGGAGCTAATTCAGAA 1254
Db      1198 ACGTAATACATGGCGGTACAGAGGTTGCGCAACCCCGAGGGGAGCTAATTCAGAA 1257
Qy      1255 AGCGGTCTGATGTCGCGATGCGAGTCTGCAACTGCACTCCGTGAAATCGGAAATGCTAGT 1314
Db      1258 AGCGGTCTGATGTCGCGATGCGAGTCTGCAACTGCACTCCGTGAAATGCTAGT 1317
Qy      1315 AATGCGGATTCAGCATGTGCGGATGAATACGTTCCCGGGTCTTGTACACACCGCCGCTCA 1374
Db      1318 AATGCGGATTCAGCATGTGCGGATGAATACGTTCCCGGGTCTTGTACACACCGCCGCTCA 1377
Qy      1375 CACCATGGAGTGGGTTTCAACAGAGCAGGTATCTTAACCTTAAGAGAGCGCTTGCCA 1434
Db      1378 CACCATGGAGTGGGTTTCAACAGAGCAGGTATCTTAACCTTAAGAGAGCGCTTGCCA 1437
Qy      1435 CGGTGAGATTCAATGACTGGGGTG 1457
Db      1438 CGGTGGGGGTCAATGACTGGGGTG 1460

RESULT 22
US-10-168-337A-10
```

```

; Sequence 10, Application US/10168337A
; Publication No. US20030170654A1
; GENERAL INFORMATION:
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholtz, Philip; and Blackall,
; APPLICANT: Linda L.
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyphosphate
; TITLE OF INVENTION: Accumulating Organisms in Wastewater
; FILE REFERENCE: 002367PC/KP
; CURRENT APPLICATION NUMBER: US/10/168,337A
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Propionibacter pelophilus
; US-10-168-337A-10

```

```

Query Match      82.8%; Score 1206.4; DB 6; Length 1485;
Best Local Similarity 90.0%; Pred. No. 2e-311;
Matches 116; Conservative 0; Mismatches 141; Indels 5; Gaps 2;

```

```

QY 1 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTGAACGGCAGACGGATGCTTGCAAT 60
DB 8 ATTGAACGCTGGCGGCGATCTTTACATGCAAGTGAACGGCAGACGGATGCTTGCAAT 67
QY 61 CTGGTGGCGAGTGGCGGAGCGGGTGAAGTATGCTCGGAACGTAATCCAGAAAGAGGGGCTA 120
DB 68 CTGATGGCGAGTGGCGGAGCGGGTGAAGTATGCTCGGAACGTAATCCAGAAAGTGGGGAT 127
QY 121 ACGCATCGAAGATGTCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGAGTGA 180
DB 128 AGTGAACGAAAGTATGCTAATACCGCATATATCTTAAGAGAGAAAGCAGGGAGTGA 187
QY 181 GACCTTGGCGTTTGAAGGCGCGATGTCTGATTACTAGTTGGTGGGTAAGGCTTAC 240
DB 188 GACCTTGGCGTTTGAAGGCGCGATGTCTGATTACTAGTTGGTGGGTAAGGCTTAC 247
QY 241 CAAGGCGAGATGATGATGTCTGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300
DB 248 CAAGGCGAGATGATGATGTCTGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 307
QY 301 GCCCAGACTCTTACCGGAGGCGAGCAGTGGGAAATTTTGAACAATGGGCGCAGCCTGATC 360
DB 308 GCCCAGACTCTTACCGGAGGCGAGCAGTGGGAAATTTTGAACAATGGGCGCAGCCTGATC 367
QY 361 CAGCAATGCCGCGTGAAGTGAAGAGCCTTGGGTTGTAAGCTCTTTCACTGAGAAAG 420
DB 368 CAGCAATGCCGCGTGAAGTGAAGAGCCTTGGGTTGTAAGCTCTTTCACTGAGAAAG 427
QY 421 AAGGTTAAGTAAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 480
DB 428 AAGGTTAAGTAAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 487
QY 481 TACGTCCAGCAGCAGCGCGTAAATACGTAGGTCAGACGCTTAATCGGAATTAAGTGGCGT 540
DB 488 TACGTCCAGCAGCAGCGCGTAAATACGTAGGTCAGACGCTTAATCGGAATTAAGTGGCGT 547
QY 541 AAGGTTGCGCAGCGCGCTTTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 600
DB 548 AAGGTTGCGCAGCGCGCTTTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 607
QY 601 TGGCTTGAATCAAGAGCTAGAGTGTGAGAGAGGAGGTAATCCATGTTAGCAG 660
DB 608 GGCCTTTGAGACTGACAGCGCTAGAGTGTGAGAGAGGAGGTAATCCATGTTAGCAG 667
QY 661 TGAATATCGTAGATATGGAAGAACATGATGCGAGGAGCAGCCTCTGGGTTAACACT 720
DB 668 TGAATATCGTAGATATGGAAGAACATGATGCGAGGAGCAGCCTCTGGGTTAACACT 727
QY 721 GAGCCTATGACGAAGAGCGTGGGAGCAACAGGATTAATACCTGTTAGTCAAGCC 780
DB 728 GAGCCTATGACGAAGAGCGTGGGAGCAACAGGATTAATACCTGTTAGTCAAGCC 787

```

```

QY 781 CTAACGATGTCATAGTGTGG--GCCTATTAGGCTTGTAACGAGCTTAACGGCT 838
DB 788 CTAACGATGTCATAGTGTGG--GCCTATTAGGCTTGTAACGAGCTTAACGGCT 847
QY 839 GAGTTGACCGCCTGGGAGTACGTCGCAAGATTAAACTCAAGAAATTGACGGGAC 898
DB 848 GAGTTGACCGCCTGGGAGTACGTCGCAAGATTAAACTCAAGAAATTGACGGGAC 907
QY 899 CCGCAAGCGGTGATATGATGATTAATTCATGATGCAAGCGGAAACCTTACTACCC 958
DB 908 CCGCAAGCGGTGATATGATGATTAATTCATGATGCAAGCGGAAACCTTACTACCC 967
QY 959 TTGACATGACGAATTTTCTAGAGATGATTAGTC---TTGGGAACGCTTAACAGG 1015
DB 968 TTGACATGACGAATTTTCTAGAGATGATTAGTC---TTGGGAACGCTTAACAGG 1027
QY 1016 TGCTCATGCTGTCTGACGCTGTCGTGAGATGTTGGGTTAAGTCCCGAACAGCG 1075
DB 1028 TGCTCATGCTGTCTGACGCTGTCGTGAGATGTTGGGTTAAGTCCCGAACAGCG 1087
QY 1076 CAACCTTGTCAATTAATGCAATCTTTGGGCACTTAATGAGACTGCGGTGACA 1135
DB 1088 CAACCTTGTCAATTAATGCAATCTTTGGGCACTTAATGAGACTGCGGTGACA 1147
QY 1136 AACCGAGAGAGTGGGATGACGTCAGTCTCTCATGGCCCTTATGGGAGGCTTCACA 1195
DB 1148 AACCGAGAGAGTGGGATGACGTCAGTCTCTCATGGCCCTTATGGGAGGCTTCACA 1207
QY 1196 CGTAATPACAATGCGCGTACAGAGGTTGCCAACCCGCGAGGGAGCTTAATCTAGAA 1255
DB 1208 CGTAATPACAATGCGCGTACAGAGGTTGCCAACCCGCGAGGGAGCTTAATCTAGAA 1267
QY 1256 GCGCGTGTAGTCCGAGATCGAGTGTGCACTCGCTGAAGTGGAAATCGCTAGTA 1315
DB 1268 GCGCGTGTAGTCCGAGATCGAGTGTGCACTCGCTGAAGTGGAAATCGCTAGTA 1327
QY 1316 ATCGGAGATCAGCATGTGCGGTGAATACGTTCCCGGGCTTGTACACAGCGCCGTAC 1375
DB 1328 ATCGGAGATCAGCATGTGCGGTGAATACGTTCCCGGGCTTGTACACAGCGCCGTAC 1387
QY 1376 ACCATGGAGGAGGTTTTCACAGAGCAGTGTCTTAACCGTAAGAGGCGCTTGCCAC 1435
DB 1388 ACCATGGAGGAGGTTTTCACAGAGCAGTGTCTTAACCGTAAGAGGCGCTTGCCAC 1447
QY 1436 GGTGAGATTCAATGACTGGGTG 1457
DB 1448 GGTGAGATTCAATGACTGGGTG 1469

```

```

RESULT 23
US-10-168-337A-6
; Sequence 6, Application US/10168337A
; Publication No. US20030170654A1
; GENERAL INFORMATION:
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholtz, Philip; and Blackall,
; APPLICANT: Linda L.
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyphosphate
; TITLE OF INVENTION: Accumulating Organisms in Wastewater
; FILE REFERENCE: 002367PC/KP
; CURRENT APPLICATION NUMBER: US/10/168,337A
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:
; OTHER INFORMATION: Polyphosphate-accumulating organism
US-10-168-337A-6

```

Query Match 81.9%; Score 1194; DB 6; Length 1460;  
 Best Local Similarity 89.9%; Pred No. 4e-308;  
 Matches 1315; Conservative 0; Mismatches 140; Indels 7; Gaps 3;

```

Oy 1 ATTGAACGCTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGCAGCGATCTTGCAAT 60
Db 1 ATTAAACGCTGGCGGCATGCTTTACATGCAAGTCGAACGGCAGCAGCGATCTTGCAAC-- 58
Oy 61 CTGGTGGAGTGGGCGGACGGGTGAGTAATGATGCGAAACGTAATCAAGAGGGGGGTA 120
Db 59 CTGGTGGAGTGGGCGGACGGGTGAGTAATGATGCGAAACGTAATCAAGAGGGGGGTA 118
Oy 121 ACGCATCGAAAGATGTGCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAAA 180
Db 119 ACGCAGCGAAAGCTACGCTAATACCGCATATCTCTGAGCAGAGAAAGCAGGGGATCGCAA 178
Oy 181 GACCTTGGCTTTTGGAGCGGCGGATGTCGATTAAGCTAAGTTGGTGGGGTAAAGGCTTAC 240
Db 179 GACCTTGGCTTTTGGAGCGGCGGATGTCGATTAAGCTAAGTTGGTGGGGTAAAGGCTTAC 238
Oy 241 CAAGCGCAGCATCAATAGTTGGTCTGAGAGAGCAGACAGCCCACTGGGAGCTGAGACACG 300
Db 239 CAAGCGCAGCATCCCTAGCGGGTCTGAGAGAGATGATCCGCCCACTGGGAGCTGAGACACG 298
Oy 301 GCCCAGACTCTCTACGGAGAGCAGCAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 360
Db 299 GCCCAGACTCTCTACGGAGAGCAGCAGTGGGGAAATTTTGGACAAATGGGCGCAAGCTGATC 358
Oy 361 CAGCAATGCCCGGCTGAGTGAAGAGAGCCTTCGGGTTGTAAAGCTCTTTCAATCGAAGAA 420
Db 359 CAGCAATGCCCGGCTGAGTGAAGAGAGCCTTCGGGTTGTAAAGCTCTTTCAAGGAGAA 418
Oy 421 AAGGTTACGGTAATTAATCGTACTCATGACGGTATGACAGAGAAAGAACCCGCTTAC 480
Db 419 AATTCCTTGGGTTAATCCCTGAGATGATGACGGTATGACCGAATTAAGAACCCGCTTAC 478
Oy 481 TACGTCAGCAGCAGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGAATTAATCTGGGGT 540
Db 479 TACGTCAGCAGCAGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGAATTAATCTGGGGT 538
Oy 541 AAGGGTGCAGCAGCGCGCTTTGTAAAGTCAGATGTAAATCCCGGGCTTAAACCTGGGAAT 600
Db 539 AAGGGTGCAGCAGCGCGCTTTGTAAAGTCAGATGTAAATCCCGGGCTTAAACCTGGGAAT 598
Oy 601 TGCCTTTGAAACTAACAAGCTAAGATGTGCGAGAGAGGTGGAATTCATGATGTAGCAG 660
Db 599 TGCATTTGAGACTGCAAGACTGAGATTTTGCAGAGAGGGGTGGAATTCACGATGAGCAG 658
Oy 661 TGAATGCGTGAAGATATGAAAGAACATCGATGGCGAAGGCGAGCTCTGGGTTAAACT 720
Db 659 TGAATGCGTGAAGATATGAGAGAACACCGATGGCGAAGGCGAGCTCTGGGTTAAACT 718
Oy 721 GACGCTCATGCAAGAACGCTGGGGAGCAACAGGATTAGATACCTGTGATGCCAGCC 780
Db 719 GACGCTCATGCAAGAACGCTGGGGAGCAACAGGATTAGATACCTGTGATGCCAGCC 778
Oy 781 CTTAAACGATGTCAACTAGTTGG--GCCTTATTAAGGCTGTGTAACGAGCTAACGGCT 838
Db 779 CTTAAACGATGTCAACTAGTTGGAGGAGTTAAACCTTTTATGATGCCGTACTCAACGGCT 838
Oy 839 GAAAGTTGACCGGCTGGGAGTACGGTCCGACAGATTAAACCTCAAGGAATTGACGGGGAC 898
Db 839 GAAAGTTGACCGGCTGGGAGTACGGTCCGACAGATTAAACCTCAAGGAATTGACGGGGAC 898
Oy 899 CCGCAGCAAGCGGTGATATGATGATTAAATTCGATGCAACGCAAAAACCTTACTTACC 958
Db 899 CCGCAGCAAGCGGTGATATGATGATTAAATTCGATGCAACGCAAAAACCTTACTTACC 958
Oy 959 TTGACATGTAGCGAAATTTTCTAGAGATAGATTAGTCT--TCGGGAACGCTTAACACAG 1015
Db 959 TTGACATGTAGCGAAATTTTCTAGAGATAGATTAGTCT--TCGGGAACGCTTAACACAG 1015
Oy 1016 TGCTGACATGGCTGTGTCAGGCTCGTGTGTAAGATGTTGGGTTAAGTCCCGCAACGAGC 1075
  
```

```

Db 1019 TGCTGACATGGCTGTGTCAGGCTCGTGTGTAAGATGTTGGGTTAAGTCCCGCAACGAGC 1078
Oy 1076 CAACCTTTGTATTAATTTGCCATCATTTGGTGGGCACTTAAATGAGACTGCCGTGACA 1135
Db 1079 CAACCTTTGTATTAATTTGCCATCATTTGAGTGGGCACTTAAATGAGACTGCCGTGACA 1138
Oy 1136 AACCGAGAAAGATGGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTGACA 1195
Db 1139 AACCGAGAAAGATGGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTGACA 1198
Oy 1196 CGTAATCAATGAGCGCGTACAGAGGTTGCCAACCCGCGAGAGGGAGCTAATCTCAGAAA 1255
Db 1199 CGTATCAATGATGTCGGTCCAGAGGTTGCCAACCCGCGAGAGGGAGCAATCTCAGAAA 1258
Oy 1256 GCGCGTGTAGTCCGATTCGAGATCTGCAACTGCAATCCGTGAAGTGGGAATCGCTAGTA 1315
Db 1259 GCCGATGTAGTCCGATTCGAGATCTGCAACTGCAATCCGTGAAGTGGGAATCGCTAGTA 1318
Oy 1316 ATCGCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCAC 1375
Db 1319 ATCGCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCAC 1378
Oy 1376 ACCATGGAGAGTGGGTTTCAACGAGAGAGTATCTAACGTTAAGAGAGGGGCTTGCAC 1435
Db 1379 ACCATGGAGAGGAGTTTGCAGAGAGTATGATGCTTAACGCAAGAGGGGATTAACAC 1438
Oy 1436 GGTAGATTCATGACTGGGGTG 1457
Db 1439 GGCAGGGTTCGTGACTGGGGTG 1460

```

RESULT 24  
 US-10-168-337A-8  
 ; Sequence 8, Application US/10168337A  
 ; Publication No. US20030170654A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholz, Philip; and Blackall,  
 ; TITLE OF INVENTION: Probes and Primers for the Detection of Polyphosphate  
 ; FILE REFERENCE: 002367PC/KF  
 ; CURRENT APPLICATION NUMBER: US/10/168,337A  
 ; CURRENT FILING DATE: 2000-12-27  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 1459  
 ; TYPE: DNA  
 ; ORGANISM: Unknown Organism  
 ; FEATURE: Description of Unknown Organism:  
 ; OTHER INFORMATION: Polyphosphate-accumulating organism  
 US-10-168-337A-8

Query Match 81.1%; Score 1181; DB 6; Length 1459;  
 Best Local Similarity 89.7%; Pred. No. 1.2e-304;  
 Matches 1302; Conservative 1; Mismatches 141; Indels 7; Gaps 3;

```

Oy 12 GCGGCATGCTTTTACACATGCAAGTCGAACGGCAGCAGGATGCTTGATCTGTGGGAG 71
Db 11 GCGGCATGCTTTTACACATGCAAGTCGAACGGCAGCAGGAGGAGAAC--CTGGGTGGGAG 68
Oy 72 TGGCGAGCGGGTGAATGATGATCGGAAGCTATCCAGAGAGGGGGGTAACGATCGAAA 131
Db 69 TGGCGAGCGGGTGAATGATGATCGGAAGCTATCTGAGTGGGGGTAACGATCGAAA 128
Oy 132 GATGTCTAATACCGCATATCTCTAAGAGAGAAAGAGGGGATTCGAAAGACCTTGCGCT 191
Db 129 GTTAGCTAATACCGCATATCTCTGAGCAGAGAAAGAGGGGATTCGAAAGACCTTGCGCT 188
Oy 192 TTTGAGAGGCGCGATGTCTGATTAAGTATGTTGGTGGGTTAAAGGCTTACCAAGCGCAGA 251
  
```

Db 189 CTGGAGCGCCGATGTCGATTAGCTAGTGTGGGTAAGGCCCTTACCAAGCGGACGA 248  
 QY 252 TCAGTAGTTGGTCTGAGAGAGACGACCACTGGAGCTGAGACAGCGGCCAGACTTC 311  
 Db 249 TCCGTAGCGGGTCTGAGAGATGATCCGCGACACTGGGACTGAGACAGCGGCCAGACTTC 308  
 QY 312 TACGGAGGACAGCACTGGGGAATTTTGGACAAATGGCGCAGACCTGATCTAGCAATGCCG 371  
 Db 309 TACGGAGGACAGCACTGGGGAATTTTGGACAAATGGCGCAGACCTGATCTAGCAATGCCG 368  
 QY 372 CGTAGTAGAAGAGCCCTTCGGGTTGTAAGCTCTTTGATGAGAGAGAAAGTTACCG 431  
 Db 369 CGTAGTAGAAGAGCCCTTCGGGTTGTAAGCTCTTTGATGAGAGAGAAATTCACCGG 428  
 QY 432 TAAATATGTCATCATGACCGTATCGACAGAGAGAGACCGGCTTAACTAGCTGCGAC 491  
 Db 429 TTAATACCTGTTGTAGATACGCTACCCGAAATAGAGAGACCGGCTTAACTAGCTGCGAC 488  
 QY 492 ACCCGCGTAAATACGTAGGGTGCAGACCGTAAATCGGAATTAATCGGCGTAAAGGTCGC 551  
 Db 489 ACCCGCGTAAATACGTAGGGTGCAGACCGTAAATCGGAATTAATCGGCGTAAAGGTCGC 548  
 QY 552 AGCGGCTTTGTAAGTCAAGTGGAAATCCCGGGCTTAACTGGGGAATGGGTTGAA 611  
 Db 549 AGCGGCTTTGTAAGTCAAGTGGAAATCCCGGGCTTAACTGGGGAATGGGTTGAA 608  
 QY 612 CTACAAAGGCTAGAGTGGGAGAGAGGAGTGAATTCATGTTGACAGTGAATGCTGA 671  
 Db 609 CTGCGAGGCTGAGGTTGGCAGAGGGGGGTGAAATTCAGCTGTAGCAGTGAATGCTGA 668  
 QY 672 GAGATATGAGAGAACTCATGAGCGAAAGGACGCTCTGGGTTAACTAGCTGACGCTCATGC 731  
 Db 669 GAGATATGAGAGAACTCATGAGCGAAAGGACGCTCTGGGTTAACTAGCTGACGCTCATGC 728  
 QY 732 ACCGAAGCGTGGGAGCAACAGAGTTAATACCTGTTAGTCCAGCCCTTAAAGCATG 791  
 Db 729 ACCGAAGCGTGGGAGCAACAGAGTTAATACCTGTTAGTCCAGCCCTTAAAGCATG 788  
 QY 792 CAATGTTGTTGG--GCCCTTATTAAGGCTGTGTAAGAGCTAACGCGTGAATGACG 849  
 Db 789 CAATGTTGTTGGAGGTTAAACCTTTTATGTCGCTGACTAACGCGTGAATGACG 848  
 QY 850 CCTGGGAGTACCGTGCAGATTTAAATCTCAAGAAATGACGGGACCCGCAAGCG 909  
 Db 849 CCTGGGAGTACCGGCGCAAGGCTTAAATCTCAAGAAATGACGGGACCCGCAAGCG 908  
 QY 910 GTGATTTATGATTTAATTCGATGCAAGCGGAAATCTTACTCTACCTTGAATGTAG 969  
 Db 909 GTGATTTATGATTTAATTCGATGCAAGCGGAAATCTTACTCTACCTTGAATGTAG 968  
 QY 970 CGAATTTCTAGATGATTAAGTCT--TCGGGAACGCTAACAGAGTGTGATGCG 1026  
 Db 969 GGAATCTGAAGATTTGGAGAGTCTGCAGAGAGACCTGAAACAGAGTGTGATGCG 1028  
 QY 1027 TGTGTCAGCTCGTGTGATGAGATGTTGGTTAAGTCCCGCAAGAGCGGACCTTGTG 1086  
 Db 1029 TGTGTCAGCTCGTGTGATGAGATGTTGGTTAAGTCCCGCAAGAGCGGACCTTGTG 1088  
 QY 1087 ATTAATTTGCAATCTTTGTTGGCACTTTAATGAGACTGCGGCTGCAAAACCGGAGAA 1146  
 Db 1089 ATTAATTTGCAATCTTTGTTGGCACTTTAATGAGACTGCGCAATGCAAAACCGGAGAA 1148  
 QY 1147 GGTGGGATGACGCTCAAGTCTCTATGAGCCCTTATGGGTAAGGCTTCAACGTAATCAAT 1206  
 Db 1149 GGTGGGATGACGCTCAAGTCTCTATGAGCCCTTATGGGTAAGGCTTCAACGTAATCAAT 1208  
 QY 1207 GGCCTGTAAGAGGAGTTGCAACCGCGAGGGGAGGCTAATCTGAGAAAGCGCTGTAG 1266  
 Db 1209 GGTGGTTCAGAGGAGTTGCAACCGCGAGGGGAGGCAATCTGAGAAAGCGATGTAG 1268  
 QY 1267 TCCGATTCAGAGTGTGCACTGACTCCGTAAGTCCGATGCTAGTAATTCGGGATCA 1326  
 Db 1269 TCCGATTCAGAGTGTGCACTGACTCCGTAAGTCCGATGCTAGTAATTCGGGATCA 1328

QY 1327 GCATGTCGGGTGAATAGTTCCGGGCTTGTACACACCGCCGTCACACCATGGAGT 1386  
 Db 1329 GCATGTCGGGTGAATAGTTCCGGGCTTGTACACACCGCCGTCACACCATGGAGT 1388  
 QY 1387 GGGTTTCACAGAGCAGTAGTGTAAACCGTAAGAGGGCGCTTGCACGGTGAATTC 1446  
 Db 1389 GGGTTTCACAGAGCAGTAGTGTAAACCGTAAGAGGGCGCTTGCACGGTGAATTC 1448  
 QY 1447 TGAATGGGGTG 1457  
 Db 1449 TGAATGGGGTG 1459  
 RESULT 25  
 US-10-168-337A-4  
 ; Sequence 4, Application US/10168337A  
 ; Publication No. US20030170654A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholz, Philip; and Blackall,  
 ; APPLICANT: Linda L.  
 ; TITLE OF INVENTION: Probes and Primers for the Detection of Polynophosphate  
 ; FILE REFERENCE: 002367PC/KF  
 ; CURRENT APPLICATION NUMBER: US/10/168,337A  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 1460  
 ; TYPE: DNA  
 ; ORGANISM: Rhodocyclus tenuis  
 ; FEATURE:  
 ; NAME/KEY: misc. feature  
 ; LOCATION: 1045, 1315  
 ; OTHER INFORMATION: n = unknown  
 US-10-168-337A-4  
 Query Match 80.9%; Score 1179.2; DB 6; Length 1460;  
 Best Local Similarity 89.3%; Pred. No. 3.5e-304;  
 Matches 1305; Conservative 0; Mismatches 150; Indels 7; Gaps 3:  
 QY 1 ATTTGAACGCTGGCGGATGCTTTTACATGCAAGTCGAAACGCGACAGCGATGCTTGCAT 60  
 Db 1 ATTTGAACGCTGGCGGATGCTTTTACATGCAAGTCGAAACGCGACAGCGATGCTTGCAT-- 58  
 QY 61 CTGGTGGCAGTGGCGGACGCGGTGATATGATCGAACTGATCCAGAAAGGGGGGATA 120  
 Db 59 CTGGTGGCAGTGGCGGACGCGGTGATATGATCGAACTGATCCAGAAAGGGGGGATA 118  
 QY 121 ACGCATGGAAGATGTCTAATACCGCATATCTCTAGAGAGAAAGCAGGGGATTCGAA 180  
 Db 119 ACGTACGGAAGTATGCTAATACCGCATATCTCTAGAGAGAAAGCAGGGGATTCGAA 178  
 QY 181 GACCTTGGCTTTTGAAGGCGCGCATGCTGATTACTGTTGGTGGGTTAAAGGCTTAC 240  
 Db 179 GACCTTGGCTTTTGAAGGCGCGCATGCTGATTACTGTTGGTGGGTTAAAGGCTTAC 238  
 QY 241 CAAGGCGAGATCAGTAGTTGTCTGAGAGAGCAACCACTGAGGATTCGAGACAG 300  
 Db 239 CAAGGCGAGATCAGTAGTTGTCTGAGAGAGCAACCACTGAGGATTCGAGACAG 298  
 QY 301 GCCCAGACTCTTACGCGGAGGACAGTGGGAAATTTTGAACAATGGGCGCAAGCTGATC 360  
 Db 299 GCCCAGACTCTTACGCGGAGGACAGTGGGAAATTTTGAACAATGGGCGCAAGCTGATC 358  
 QY 361 CAGCATGCGCGCTGAGTGAAGAGGCTTCGGGTTGTAAGGCTCTTCAAGTGAAGA 420  
 Db 359 CAGCATGCGCGCTGAGTGAAGAGGCTTCGGGTTGTAAGGCTCTTCAAGTGAAGA 418  
 QY 421 AAGGTTACGCTAATATCTGATCATGACGCTATCGACAGAAAGCAACCGGCTAAC 480  
 Db 419 AATTTGCTCAGAGTAATACCTGAGTAGATGACGCTACCGAAGAAAGCAACCGGCTAAC 478

```
QY 481 TACGTCGACGAGCCGCGGTAACTAGTAGGGTGAAGCGTTAATCGAATTACTGGGGCT 540
Db 479 TACGTCGACGAGCCGCGGTAACTAGTAGGGTGAAGCGTTAATCGAATTACTGGGGCT 538
QY 541 AAAAGGTGGCAGCGCGCTTTGTAAATCATGATGTGAATCCCGGGCTTAACTTGGGAAT 600
Db 539 AAAAGGTGGCAGCGCGCTTTGTAAATCATGATGTGAATCCCGGGCTTAACTTGGGAAT 598
QY 601 TCGCTTTGAAATCTACAGGCTAGATGTGTGGCAGAGGAGTGTGAATTCATGTGTAGCAG 660
Db 599 TCGCTTTGAGCTGACGACTAGAGTGTGTGGCAGAGGAGTGTGAATTCATGTGTAGCAG 658
QY 661 TGAATTCGCTAGATATATGGAAGAACTATGATGGGAGGAGCGCTCCGCGGTAACT 720
Db 659 TGAATTCGCTAGATATGGAAGAACTATGATGGGAGGAGCGCTCCGCGGTAACT 718
QY 721 GACGCTCATGACGAAAGCGTGGGAGGACAAACAGATTAGATACCTGGTAGTCCAGCGC 780
Db 719 GACGCTCATGACGAAAGCGTGGGAGGACAAACAGATTAGATACCTGGTAGTCCAGCGC 778
QY 781 CTAAACGATGTCACTAGTTGTTG--GGCTTATTAGGCTTGTGTAACGAACTAACGCGT 838
Db 779 CTAAACGATGTCACTAGTTGTTGTTGGGTGTTAAACCTATTAGTCCGCTAGCTAACGCGT 838
QY 839 GAAGTTGACCGGCTGGGAGTACGGTCCGCAAGATTAAACCTGAAGAAATTGACGGGGAC 898
Db 839 GAAGTTGACCGGCTGGGAGTACGGTCCGCAAGATTAAACCTGAAGAAATTGACGGGGAC 898
QY 899 CCGCAAGACGGTGGATTATGTGATTTAATTCGATGCAACCGCAAAAACCTTACTTACC 958
Db 899 CCGCAAGACGGTGGATTATGTGATTTAATTCGATGCAACCGCAAAAACCTTACTTACC 958
QY 959 TTGAATGATGAGGAATTTTCTAGAGATAGATTAGTGC--TTCCGGAAACGCTAACACAG 1015
Db 959 TTGAATGATGAGGAATTTTCTAGAGATAGATTAGTGC--TTCCGGAAACGCTAACACAG 1018
QY 1016 TGCTGCATGGGCTGTGTGTCAGCTCGTGTGATGTTGGGTTAAGTCCCGCAAGAGAG 1075
Db 1019 TGCTGCATGGGCTGTGTGTCAGCTCGTGTGATGTTGGGTTAAGTCCCGCAAGAGAG 1078
QY 1076 CAACCTTGTCTAATTAATTCATTTGGTGGGCACTTTAATGAGATCGCGGTGACA 1135
Db 1079 CAACCTTGTCTAATTAATTCATTTAGTTGGGCACTTTAATGAGATCGCGGTGACA 1138
QY 1136 AACCGAGAGAGTGGGATGACGTCAAGTCTTATGCGCTTATGGGTAGGGCTTACA 1195
Db 1139 AACCGAGAGAGTGGGATGACGTCAAGTCTTATGCGCTTATGGGTAGGGCTTACA 1198
QY 1196 CGTATTAAGTGGCGCGCTACAGAGGTTGCCAACCGGAGGGGAGGCTAATCTCAGAAA 1255
Db 1199 CGTATTAAGTGGCGCGCTACAGAGGTTGCCAACCGGAGGGGAGGCTAATCTCAGAAA 1258
QY 1256 GCGCGTCTAGTCCGGATCGAGTCTGCAACTGCACTCCGTGAAGTCCGAAATCGCTAGTA 1315
Db 1259 GCGCGTCTAGTCCGGATCGAGTCTGCAACTGCACTCCGTGAAGTCCGAAATCGCTAGTA 1318
QY 1316 ATTCGGGATCAGCATGTGCGGGTGAATAGTTCCCGGGCTTTGTACACACCGCCCTGAC 1375
Db 1319 ATTCGGGATCAGCATGTGCGGGTGAATAGTTCCCGGGCTTTGTACACACCGCCCTGAC 1378
QY 1376 AACCATGGAGTGGGTTTACCAAGAGCAGTGTGAATCCGTAAGAGAGGGCGCTTGACAC 1435
Db 1379 AACCATGGAGTGGGTTTACCAAGAGTGTGAATCCGTAAGAGAGGGCGATTAACAC 1438
QY 1436 GGTGATTCATGACTGGGGTG 1457
Db 1439 GGTGATTCATGACTGGGGTG 1460
```

```
Publication No. US20030170654A1
GENERAL INFORMATION:
APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholtz, Philip; and Blackall,
APPLICANT: Linda L.
TITLE OF INVENTION: Probes and Primers for the Detection of Polynophosphate
FILE REFERENCE: 002367PC/KF
CURRENT APPLICATION NUMBER: US/10/168,337A
CURRENT FILING DATE: 2000-12-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1460
TYPE: DNA
ORGANISM: Rhodocyclus tenuis
FEATURE:
NAME/KEY: misc.feature
LOCATION: 50, 59
OTHER INFORMATION: n = unknown
US-10-168-337A-2
```

Query Match 79.4%; Score 1156.8; DB 6; Length 1460;  
Best Local Similarity 88.3%; Pred. No. 3.3e-258;  
Matches 1291; Conservative 0; Mismatches 164; Indels 7; Gaps 3;

```
QY 1 ATGAAAGCTGGCGGCAATGCTTTACATGCAAGTCGACGCGACGAGTCTTGCAT 60
Db 1 ATGAAAGCTGGCGGCAATGCTTTACATGCAAGTCGACGCGACGAGTCTTGCAT 58
QY 61 CTGTGGCGAGTGGCGGACGGGTGATATGATCGAAGCTATCCAGAAAGGGGGATA 120
Db 59 NTGGCGACAGATGGCGAAGCGGTGATATGATCGAAGCTATCCAGAAAGGGGGATA 118
QY 121 ACGATGGAAGATGTCTATATACCGATATATCTTAAAGAGAAACACAGGGATTCGAAA 180
Db 119 ACGTGGGAAAGTTACCTTAATACCGATATCTTAAAGAGAAACACAGGGATTCGAAA 178
QY 181 GACCTTGCGCTTTTGGAGCGCGCGATGTCTGATTAAGTATGTTGGGTAAAGGCTTAC 240
Db 179 GACCTTGCGCTTTTGGAGCGCGCGATGTCTGATTAAGTATGTTGGGTAAAGGCTTAC 238
QY 241 CAAGGCGACATCGATAGTGTCTGAGAGAGCAGCAGCCACACTGGGACTGAGACAG 300
Db 239 CAAGGCGACATCGATAGTGTCTGAGAGAGCAGCAGCCACACTGGGACTGAGACAG 298
QY 301 GCCCAGACTCTTACGGGAGCAGCAGTGGGGAATTTTGGACATGGGCGCAAGCTGATC 360
Db 299 GCCCAGACTCTTACGGGAGCAGCAGTGGGGAATTTTGGACATGGGCGCAAGCTGATC 358
QY 361 CAGCAATGGCGCGGAGTGAAGAAAGGCTTCCGGTTTAAAGCTCTTCAGTGCAAGAA 420
Db 359 CAGCAATGGCGCGGAGTGAAGAAAGGCTTCCGGTTTAAAGCTCTTCAGTGCAAGAA 418
QY 421 AAAGTTACGGTAAATTAATCGTGAATCGTATCGAAGGTAACAGAACAGCAGGCTTAC 480
Db 419 AAAGTTACGGTAAATTAATCGTGAATCGTATCGAAGGTAACAGAACAGCAGGCTTAC 478
QY 481 TACGTCGACAGACCGCGGTATATCGTAGGGTGAACGCTTAACTCGAATTAATCGGGCT 540
Db 479 TACGTCGACAGACCGCGGTATATCGTAGGGTGAACGCTTAACTCGAATTAATCGGGCT 538
QY 541 AAAAGGTGGCAGCGCGCTTTGTAAATCATGATGTGAATCCCGGGCTTAACTTGGGAAT 600
Db 539 AAAAGGTGGCAGCGCGCTTTGTAAATCATGATGTGAATCCCGGGCTTAACTTGGGAAT 598
QY 601 TCGCTTTGAAATCTACAGGCTAGATGTGTGGCAGAGGAGTGTGAATTCATGTGTAGCAG 660
Db 599 TCGCTTTGTAGCTGACAGCTAGAGTGTGCAAGAGGGGGTGGAAATTCACGTTAGCAG 658
QY 661 TGAATTCGCTAGATATATGGAAGAACTATGATGGGAGGAGCGCTCCGCGGTAACT 720
Db 659 TGAATTCGCTAGATATATGGAAGAACTATGATGGGAGGAGCGCTCCGCGGTAACT 718
```

```

QY 721 GAGCCTCATGCAGAAAGCGTGGGAGCAAAACAGATTTAGATACCTCGTAGTCACAGCC 780
Db 719 GAGCCTCATGCAGAAAGCGTGGGAGCAAAACAGATTTAGATACCTCGTAGTCACAGCC 778
QY 781 CTAAACGATGCACATGATTGTTG--GGCCCTTTAGAGCTTGGTAACGAAGCTAACGCGT 838
Db 779 CTAAACGATGCACATGATTGTTGTTGGGTTAAACCAATTAGTCCGTAGCTAACCGCT 838
QY 839 GAAGTTGACCGCCTGGGAGATACCGTGCAGAAATTAATACTCAAGGAATTGACGGGAGC 898
Db 839 GAATTTGACCGCCTGGGAGATACCGCGCAAGGTTAAATCTCAAGGAATTGACGGGAGC 898
QY 899 CCGCAACAAGCGGTGGATTAATGTGGATTAATTCATGCAAGCGCAAAAACCTTACCTTACC 958
Db 899 CCGCAACAAGCGGTGGATTAATGTGGATTAATTCATGCAAGCGCAAAAACCTTACCTTACC 958
QY 959 TTGACATGTACGGAATTTTCTAGAGATAGATTAGTGC---TTGGGAACGCTAACAGAG 1015
Db 959 TTGACATGTACGGAATCTTGAAGATTAAGGAAGTGCAGGAAGGAACTGAACACAGG 1018
QY 1016 TGCTGCATAGCTGTCGTACAGCTCGTGTCTGAGATGTTGGTTTAAGTCCGCAACGAGCG 1075
Db 1019 TGCTGCATAGCTGTCGTACAGCTCGTGTCTGAGATGTTGGTTTAAGTCCGCAACGAGCG 1078
QY 1076 CAACCTTGTCTAATTAATTTGCCATCATATTGGTTGGGCACTTTAATGAACATGCGGTGACA 1135
Db 1079 CAACCTTGTCTAATTAATTTGCCATCATATTGGGCACTCTTAATGAACATGCGGTGACA 1138
QY 1136 AACCGGAGGAAGTGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTGACA 1195
Db 1139 AACCGGAGGAAGTGGGGATGACGTCAAGTCTCATGAGCCCTTAATGGGTAGGGCTTGACA 1198
QY 1196 CGTAATACATATGCGCGCTGACAGAGGGTTGCCAACCCCGAGGGGAGCTAATCTCAGAAA 1255
Db 1199 CGTATACATATGTCGTGTCGTCAGAGGGTTGCCAACCCCGAGGGGAGCTAATCTCCGCAAA 1258
QY 1256 GCGCGCTGTAGTCCGAGATCGAGATCTGCAACTCGACTCCGTGAAGTGGGATTCGTGACTA 1315
Db 1259 GCGCATGTAGTCCGAGATTTGGCATGCTGCMACTGCACTGCAATGAAGTCCGAATTCCTGACTA 1318
QY 1316 ATCGGGGATCAGCAGATGTCGCGGTGAATACGTTCCCGGGCTTTGTACACACCGCCGATCAC 1375
Db 1319 ATCGGGGATCAGCAGATGTCGCGGTGAATACGTTCCCGGGCTTTGTACACACCGCCGATCAC 1378
QY 1376 ACCATGGAGTGGGTTTTCACACGAAAGCAGTAGTCTAAACGTTAAGAGAGGCGCTTGCCAC 1435
Db 1379 ACCATGGAGTGGGTTTTCGCCAGAGTAGTACTTAACCGAAGAGAGGCGGATTAACAC 1438
QY 1436 GGTGAGATTGACTGAGGCGGTG 1457
Db 1439 GGCAGGGTTGCTGACTGGGCGTG 1460

RESULT 27
US-10-168-337A-1
; Sequence 1, Application US/10168337A
; Publication No. US20030170654A1
; GENERAL INFORMATION:
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholz, Philip; and Blackall,
; APPLICANT: Linda L.
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyphosphate
; TITLE OF INVENTION: Accumulating Organisms in Wastewater
; FILE REFERENCE: 002367PC/KF
; CURRENT APPLICATION NUMBER: US/10/168,337A
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Rhodocyclus tenuis
; FEATURE:
; NAME/KEY: misc_feature

```

Query Match	Similarity	88.2%	Pred. No. 6,6e-298;	Mismatches 165;	Indels 7;	Gaps 3;
Query Match	Similarity	88.2%	Pred. No. 6,6e-298;	Mismatches 165;	Indels 7;	Gaps 3;
Best Local	Similarity	88.2%	Pred. No. 6,6e-298;	Mismatches 165;	Indels 7;	Gaps 3;
Matches 1290;	Conservative	0;	Mismatches 165;	Indels 7;	Gaps 3;	
US-10-168-337A-1						
LOCATION: 50, 59, 881						
OTHER INFORMATION: n = unknown, m = a or c						
79.3%;	Score 1155.8;	DB 6;	Length 1460;			
1	ATTGAGCGCTGGGGGCAATGCTTTTACATATCAGTACGCAACGGGACACGAGCTCTTGAT	60				
1	ATTGAGCGCTGGGGGCAATGCTTTTACATATCAGTACGCAACGGGACACGAGCTCTTGAT	58				
61	CTGGTGGCGAGTGGCGGACCGGCTGATGATGATCGGACGATTCGAAAGAGCGGGGCTA	120				
59	NTGGCGACAGATGGCGGAAACGGGTGATGATGATCGGACGAGCTGAGCTGAGTGGGGATTA	118				
121	ACGCATCGAAGATGCTGCTAATCCGCAATATCTTAAGAGGAAACAGGGGATCGAA	180				
119	ACGTAGCGAAGATTACGCTAATACCGCATATCTGTGAGACGAAACAGGGGATCTTAG	178				
181	GACCTTGGGCTTTTGGAGCGGCGGATGCTGATAGTAGTGGTGGGCTAAAGGCGCTAC	240				
179	GACCTTGGGCTTTTGGAGCGGCGGATGCTGATAGTAGTGGTGGGCTAAAGGCGCTAC	238				
241	CAAGCGACGATTCAGTAGTGTGTGAGAGGACGACGACCACTGGGACTGAGACG	300				
239	CAAGCGACGATTCGTTAGCGGCTCTGAGAGGATGATCCGCACTGGGACTGAGACG	298				
301	GCCGACAGCTCTTAACGGGAGGACGACGATGGGGAAATTTTGGACATGGGCGCAGCTGATC	360				
299	GCCGACAGCTCTTAACGGGAGGACGACGATGGGGAAATTTTGGACATGGGCGGAAACCTGATC	358				
361	CAGCAATCCCGCTGAGAGGAAAGGCGCTTGGGGTGTAAAGCTCTTTCAGTCCAGAGA	420				
359	CAGCATCCCGCTGAGAGGAAAGGCGCTTGGGGTGTAAAGCTCTTTCAGTCCAGAGA	418				
421	AAAGGTTACGGTAATTAATCGTGACTCATGACGGTATGACAGAGAAAGCACCGGCTAAC	480				
419	AATGMACTGCTAATTAATCGTGACTCATGACGGTATGACAGAGAAAGCACCGGCTAAC	478				
481	TACGTGCCAGACCGCGGTAATACGTAAGGCTGCAAGCTTAATTCGAAATTAATCTGGCGT	540				
479	TACGTGCCAGACCGCGGTAATACGTAAGGCTGCAAGCTTAATTCGAAATTAATCTGGCGT	538				
541	AAAGGTGGCGAGCGGCTTTGTAAGTACGATGGAATCCCGGGCTTAACCTGGGAAT	600				
539	AAAGGTGGCGAGCGGCTTTGTAAGTACGATGGAATCCCGGGCTTAACCTGGGAAT	598				
601	TGCGTTTGAATCTAACAGGCTTACGATGTGGGAGAGGAGTGGAAATTCATGTGTAGAG	660				
599	TGCGTTTGAATCTAACAGGCTTACGATGTGGGAGAGGAGTGGAAATTCATGTGTAGAG	658				
661	TGAAATGCTAGATATGGAAGAACATCGATGGCGAAGGACGCTCTGGGTTAACACT	720				
659	TGAAATGCTAGATATGGAAGAACATCGATGGCGAAGGACGCTCTGGGTTAACACT	718				
721	GACGCTCATGACGAAAGCTGGGAGCAACAGATTAAGTACCTGTGTGTCCAGGCG	780				
719	GACGCTCATGACGAAAGCTGGGAGCAACAGATTAAGTACCTGTGTGTCCAGGCG	778				
781	CTAAAGATGATCAACATGATGTTTG--GGCTTATTAAGGCTTGTATCAAGAGCTAACGCGT	838				
779	CTAAAGATGATCAACATGATGTTTG--GGCTTATTAAGGCTTGTATCAAGAGCTAACGCGT	838				
839	GAAATTGACCGCTGGGAGATACGCTCGCAGATTAACTCAAGAAATGACGGGAC	898				
839	GAAATTGACCGCTGGGAGATACGCTCGCAGATTAACTCAAGAAATGACGGGAC	898				
899	CGGACACAGCGGTGATTAATGTGATTAATTCGATGCAACGGGAAACCTTAACCTACC	958				
899	CGGACACACAGCGGTGATTAATGTGATTAATTCGATGCAACGGGAAACCTTAACCTACC	958				
959	TTGACATGTAGGCAATTTCTAGAGATGATGATGTC--TTTCGGGAAACGCTAACAGG	1015E				





QY 1318 CGCGATCAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTAACACCCGCCCTTCAAC 1377  
DB 1342 CGCGATCAGCATGTCGCGGTGAATACGTTCCCGGGCTTTGTAACACCCGCCCTTCAAC 1401  
QY 1378 CATGGAGTGGGTTTCAACGAAGAGGATGCTAACCGGTAGGAGGGCGCTTGCCAGG 1437  
DB 1402 CATGGAGTGGGTTTCAACGAAGAGGATGCTAACCGGAAGGGGGCGATTACACCG 1461  
QY 1438 TGAGATTGATGACTGGGGTG 1457  
DB 1462 TAGGATTGATGACTGGGGTG 1481

RESULT 29  
US-10-723-365B-32  
Sequence 32, Application US/10723365B  
Publication No. US20050112590A1  
GENERAL INFORMATION:  
APPLICANT: VAN DEN BOOM, DIRK  
APPLICANT: BOCKER, SEBASTIAN  
TITLE OF INVENTION: VARIATION DETECTION AND DISCOVERY  
FILE REFERENCE: SEQ-2073-UT  
CURRENT APPLICATION NUMBER: US/10/723,365B  
PRIORITY FILING DATE: 2003-11-26  
PRIORITY APPLICATION NUMBER: 60/429,895  
PRIORITY FILING DATE: 2002-11-27  
NUMBER OF SEQ ID NOS: 85  
SOFTWARE: PatentIn Ver. 3.2  
SEQ ID NO 32  
LENGTH: 1496  
TYPE: DNA  
ORGANISM: *Bordetella pertussis*  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (821)  
OTHER INFORMATION: a, c, g, or t  
US-10-723-365B-32

Query Match 78.5%; Score 1143.8; DB 9; Length 1496;  
Best Local Similarity 88.1%; Pred. No. 9,9e-295;  
Matches 1281; Conservative 0; Mismatches 163; Indels 10; Gaps 3;

QY 7 CGCTGGCGGATGCTTTACATGCAAGTCGAAACGCGACGCGATGCTTGCATCTGTG 66  
DB 1 CGCTTACGCGGATGCTTTACATGCAAGTCGAAACGCGACGCGG--CTTGGTCTGGG 58  
QY 67 GCGAGTGGCGGACCGGTGAGTAATGCAATCGGAACGTAATCAAGAGAGGGGGTACGAT 126  
DB 59 GCGAGTGGCGGACCGGTGAGTAATGTAATCGGAACGTCGCCAGTACGCGGGATTAATACG 118  
QY 127 CGAAAGATGCTTAATCCGCTTAATCTTAAGAGAGAAACGAGGGATGAAAGACCTT 186  
DB 119 CGAAAGCTTAATGCTTAATCCGCTTAATCGGCGGAAAGCGGGGACCTTGGGGCTTC 178  
QY 187 GCGCTTTGAGAGGCGCGATGCTGATTAAGTGAATGATGGTGAAGGCTTCAAGGC 246  
DB 179 GCACTATTGAGGCGCGATGCTGATTAAGTGAATGATGGTGAAGGCTTCAAGGC 238  
QY 247 GAGCATCAGTATGTTGCTGAGAGAGACCAACGACACTGGGACTGAGACACGCGCCAG 306  
DB 239 GAGCATCAGTATGTTGCTGAGAGAGACCAACGACACTGGGACTGAGACACGCGCCAG 298  
QY 307 ACTCTTACCGGAGGACGATGCGGAAATTTTGAACATGCGCGGACCTGATCAAGCA 366  
DB 299 ACTCTTACCGGAGGACGATGCGGAAATTTTGAACATGCGCGGACCTGATCAAGCA 358  
QY 367 TGCGCGGTGAGTGAAGAGGCTTGGGGTTGTAAGCTCTTTCAGTCGAGAGAAAGT 426  
DB 359 TCCGCGGTGAGTGAAGAGGCTTGGGGTTGTAAGCTCTTTCAGTCGAGAGAAAGT 418  
QY 427 TACGTAATTAATGCTGATCATGACGATATGACAGAGAAAGACCGGCTTAATCACTG 486

DB 419 TCTGGCTAATATCTGGGCAACTGACGGTACCTGACAGAAATGAGCACCGGCTTAATCACTG 478  
QY 487 CCAGCAGCGCGGATTAATCTGAGGGTGCACACGTTAATCGAAATTAATCTGGCGTAAAGG 546  
DB 479 CCAGCAGCGCGGATTAATCTGAGGGTGCACACGTTAATCGAAATTAATCTGGCGTAAAGG 538  
QY 547 TGCGCAGCGGCTTGTAAATGATGTAATCCCGGGCTTAACTGGGAATTCGCTT 606  
DB 539 TGCGCAGCGGCTTGTAAATGATGTAATCCCGGGCTTAACTGGGAATTCGCTT 598  
QY 607 TGAACCTAACAGGCTGAGTGTGCGAGAGAGGTGAATTCATGTGTAGCATGGAAT 666  
DB 599 TTTAATCTACCGGGCTAGAGTGTGTGAGAGAGGTGAATTCGCGGTGTAGCATGGAAT 658  
QY 667 GCGTGAAGATTAATGAGAAATCCATGTCGCGAAGGACGCTCCTGGGTTAATCACTGACGT 726  
DB 659 GCGTGAAGATTAATGAGAAATCCATGTCGCGAAGGACGCTCCTGGGTTAATCACTGACGT 718  
QY 727 CATGACGAAAGCGTGGGAGAGCAACAGATTAATGATCCCTGTAGTCCACGCTTAAC 786  
DB 719 CATGACGAAAGCGTGGGAGAGCAACAGATTAATGATCCCTGTAGTCCACGCTTAAC 778  
QY 787 GATGCTAATGATGTTGGGCTTATTAAGCTTGTGAACGAAGCTTAACGCTGAAGTTGA 846  
DB 779 GATGCTAATGATGTTGGGCTTATTAAGCTTGTGAACGAAGCTTAACGCTGAAGTTGA 838  
QY 847 CGCGCTGGGAGTACGCTGCGCAAGTAAATCAAGAAATTAAGAGGAGGAGCCGCAACA 906  
DB 839 CGCGCTGGGAGTACGCTGCGCAAGTAAATCAAGAAATTAAGAGGAGGAGCCGCAACA 898  
QY 907 GCGGTGATTAATGATGTTAATTCATGATCAACGCAAAACCTTAACCTTGTACATG 966  
DB 899 GCGGTGATTAATGATGTTAATTCATGATCAACGCAAAACCTTAACCTTGTACATG 958  
QY 967 TAGCGAATTTTGAAGATAGATTAGTCT--TCGGAAACGCTTAACAGGTGCTGAT 1023  
DB 959 TCTGGAATTCGCAAGAGATTTGGCAGTGTGCGCAAGAAACGGAACAGAGTGTGAT 1018  
QY 1024 GCGTGTGCTGACCTGCTGCTGAGATGTTGGGTTAATGTCGCGCAACGAGCCCAACCTT 1083  
DB 1019 GCGTGTGCTGACCTGCTGCTGAGATGTTGGGTTAATGTCGCGCAACGAGCCCAACCTT 1078  
QY 1084 GTCATTAATTTGCTATCATTTGGTGGGCACTTAATGAGATGCGGGTGAACAACCGGAG 1143  
DB 1079 GTCATTAATTTGCTATCAAA--GGGCACTTAATGAGATGCGGGTGAACAACCGGAG 1133  
QY 1144 GAAAGTGGGATGACGTCAGATCTTCATGAGCCCTTAATGGGTAGGGCTTCAACGTAATAC 1203  
DB 1134 GAAAGTGGGATGACGTCAGATCTTCATGAGCCCTTAATGGGTAGGGCTTCAACGTAATAC 1193  
QY 1204 AATGCGCGTACAGAGGTTGCCCAACCGCGAGGGGAGCTTAATCAAGAAACCGGCTG 1263  
DB 1194 AATGCGCGGATGACAGAGGTTGCCCAACCGCGAGGGGAGCTTAATCAAGAAACCGGCTG 1253  
QY 1264 TAGTCCGATCGAGATCTGCAACTGATCCGTAATGAGATGCGTAATGCGGGA 1323  
DB 1254 TAGTCCGATCGAGATCTGCAACTGATCCGTAATGAGATGCGTAATGCGGGA 1313  
QY 1324 TCAGCATGTCGCGGTAATACGTTCCGGGCTTGTACACCGCGCTTCAACCATATGG 1383  
DB 1314 TCAGCATGTCGCGGTAATACGTTCCGGGCTTGTACACCGCGCTTCAACCATATGG 1373  
QY 1384 AGTGGTTTCAACGAGAGGATGATCTTAACGTAAGAGAGGCGCTTCCACGCTGAGAT 1443  
DB 1374 AGTGGTTTCAACGAGAGGATGATCTTAACGTAAGAGAGGCGCTTCCACGCTGAGAT 1433  
QY 1444 TCAATGCTGGGGTG 1457  
DB 1434 TCAATGCTGGGGTG 1447

RESULT 30

US-10-168-337A-9  
; Sequence 9, Application US/10168337A  
; Publication No. US20030170654A1  
; GENERAL INFORMATION:  
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholtz, Philip; and Blackall,  
; APPLICANT: Linda L.  
; TITLE OF INVENTION: Probes and Primers for the Detection of Polynophosphate  
; TITLE OF INVENTION: Accumulating Organisms in Wastewater  
; FILE REFERENCE: 002367PC/KF  
; CURRENT APPLICATION NUMBER: US/10/168.337A  
; CURRENT FILING DATE: 2000-12-27  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 1426  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:  
; OTHER INFORMATION: Polynophosphate-accumulating organism  
US-10-168-337A-9

Query Match 78.4%; Score 1142; DB 6; Length 1426;  
Best Local Similarity 89.3%; Pred. No. 3e-294;  
Matches 1266; Conservative 0; Mismatches 145; Indels 7; Gaps 3;

QY 12 GCGGCATCTTTACATGCAAGTGAACGCGACGACGCGATCTTGATCTGGTGGCAG 71  
DB 11 GCGGCATCTTTACATGCAAGTGAACGCGACGACGCGGCGCAAC--CTGGTGGCAG 68  
QY 72 TGGCGGACGGGTGATTAATGATCGGAACGTATCCAGAGAGGGGGGTAAACATCGAA 131  
DB 69 TGGCGGACGGGTGATTAATGATCGGAACGTATCCAGAGAGGGGGGTAAACATCGAA 128  
QY 132 GATGCTAATACCGCATATCTTAAGAGAGAAAGAGGGGATCGAAAGCTTGCCT 191  
DB 129 GTTAGCTAATACCGCATATCTTGTAGAGAGAAAGAGGGGATCGAAAGCTTGCCT 188  
QY 192 TTTGGAGCGCGCATGTCTGATTAGCTAGTTGGTGGGTAAAGGCTTACCAAGCGACGA 251  
DB 189 CGAGGAGCGCGCATGTCTGATTAGCTAGTTGGTGGGTAAAGGCTTACCAAGCGACGA 248  
QY 252 TCAGTAGTTGTCTGAGAGAGACGACGACCACTGGGATCTGAGACAGCGCCCACTCC 311  
DB 249 TCCGTAAGCGGTCTGAGAGAGATCCGCACTGGAACCTGAGACAGCGGTCCAGACTCC 308  
QY 312 TACGGGAGGACGACGATGGGGAAATTTTGAACAATGGGCGCAAGCTGATCCAGCAATGCC 371  
DB 309 TACGGGAGGACGACGATGGGGAAATTTTGAACAATGGGCGCAAGCTGATCCAGCAATGCC 368  
QY 372 CGTAGTAGAAGAGCGCTTCGGGTTGAAGCTCTTTGAGTGCAGAGAGAAAGTTACGG 431  
DB 369 CGTAGTAGAAGAGCGCTTCGGGTTGAAGCTCTTTGAGTGCAGAGAGAAATCGACGAG 428  
QY 432 TAAATATCTGACTCATGACGCTATCCAGAGAAAGACCGGCTTAATCTAGTCCAGC 491  
DB 429 TAAATATCTGACTCATGACGCTATCCAGAGAAAGACCGGCTTAATCTAGTCCAGC 488  
QY 492 AGCGCGGTAATAGCTAGAGGTGCAAGCGTTAATCGAATTAATCTGGGCGTAAAGGGTGGC 551  
DB 489 AGCGCGGTAATAGCTAGAGGTGCAAGCGTTAATCGAATTAATCTGGGCGTAAAGGGTGGC 548  
QY 552 AGCGCGCTTTTGAATGATGATGTAATCCCGGCGTTAACTCTGGGAAATGGGTTGAA 611  
DB 549 AGCGCGCTTTTGAATGATGATGTAATCCCGGCGTTAACTCTGGGAAATGGGTTGAA 608  
QY 612 CTACAGGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTGACGATGAATGCGTA 671  
DB 609 CTGCGCAAGCTGAGTGTGCGAGAGGAGGTGAATTCATGTGTGACGATGAATGCGTA 668  
QY 672 GAGATATGAGAAATCATGATGCGAGAGGAGCTCTCGGGGTTAACAATGACGCTCAGGC 731  
DB 669 GAGATATGAGAAATCATGATGCGAGAGGAGCTCTCGGGGTTAACAATGACGCTCAGGC 728

QY 732 ACGAAGCGTGGGAGAGCAACAGGATTAAGATACCTGTGTAGTCCACGCCCTTAAACGATGT 791  
DB 729 ACGAAGCGTGGGAGAGCAACAGGATTAAGATACCTGTGTAGTCCACGCCCTTAAACGATGT 788  
QY 792 CAACTAGTGTGTGG--GCTTATTAGCTTGTGTAACGAGCTAAACGCGTGAATTTGACCG 849  
DB 789 CAACTAGTGTGTGGAGGGTTAAACCTTTTATGTGCGGAGCTAACGCGTGAATTTGACCG 848  
QY 850 CTTGGGAGTACGCTCGCAAGATTAAACCTAAAGAAATTTAGCGGGAGCCGCGCAACGCG 909  
DB 849 CTTGGGAGTACGCTCGCGCAAGATTAAACCTAAAGAAATTTAGCGGGAGCCGCGCAACGCG 908  
QY 910 GTGATTAATGATTAATTTGATGCAACGCGGAAACCTTAACCTTACCTTGAACATGTAG 969  
DB 909 GTGATTAATGATTAATTTGATGCAACGCGGAAACCTTAACCTTGAACATGTAG 968  
QY 970 CGAATTTTCTAGATAGATTAGTGT--TCGGGAAAGCTTAACAGAGTGTCTCATGCGC 1026  
DB 969 CGAATTTTCTAGATAGATTAGTGT--TCGGGAAAGCTTAACAGAGTGTCTCATGCGC 1028  
QY 1027 TGTGCTAGCTGTGTGCGATGATGTTAGTTCCTCGCAACGAGCGCAACCTTGTCTC 1086  
DB 1029 TGTGCTAGCTGTGTGCGATGATGTTAGTTCCTCGCAACGAGCGCAACCTTGTCTC 1088  
QY 1087 ATTAATTTCCATCATTTTGTGTGGGCACTTTAATGAGATGCGCGGTGAACAAACCGGAGAA 1146  
DB 1089 ATTAATTTCCATCATTTGATGAGTGGGCACTTTAATGAGATGCGCGGTGAACAAACCGGAGAA 1148  
QY 1147 GGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGGGATGAGGCTTACACAGCTAATACAT 1206  
DB 1149 GGTGGGAGTGAAGTCAAGTCTCATGCGCTTATGGGATGAGGCTTACACAGCTAATACAT 1208  
QY 1207 GCGCGTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAAGAAAGCGCTGTAG 1266  
DB 1209 GGTGCTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAAGAAAGCGCTGTAG 1268  
QY 1267 TCCGATTCGAGTGTGCAACCTCGACTCGTGAAGTCCGAATCGTATGATTCGGGATCA 1326  
DB 1269 TCCGATTCGAGTGTGCAACCTCGACTCGTGAAGTCCGAATCGTATGATTCGGGATCA 1328  
QY 1327 GCATGTCGCGGTGAATAGTTCGCGGCTTGTACACAGCCCGTCAACATGAGAGT 1386  
DB 1329 GCATGTCGCGGTGAATAGTTCGCGGCTTGTACACAGCCCGTCAACATGAGAGT 1388  
QY 1387 GGGTTTCAACGAGAGGATGTTAACCCTGAAGAG 1424  
DB 1389 GGGTTTCAACGAGAGGATGTTAACCCTGAAGAG 1426

RESULT 31  
US-10-029-397A-41  
; Sequence 41, Application US/10029397A  
; Publication No. US20030175709A1  
; GENERAL INFORMATION:  
; APPLICANT: MURPHY, GEORGE L.  
; APPLICANT: WHITLEY, J. PENN  
; TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS  
; FILE REFERENCE: AMBI.076US  
; CURRENT APPLICATION NUMBER: US/10/029.397A  
; CURRENT FILING DATE: 2002-03-19  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 41  
; LENGTH: 1535  
; TYPE: DNA  
; ORGANISM: Burkholderia cepacia  
US-10-029-397A-41

Query Match 78.1%; Score 1138.4; DB 6; Length 1535;  
Best Local Similarity 87.7%; Pred. No. 2.7e-293;  
Matches 1280; Conservative 0; Mismatches 171; Indels 9; Gaps 3;

```

QY 1 ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGCGACACGAGATGCTTGCAT 60
DB 29 ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGCGACACGAGATGCTTGCAT 88
QY 61 CTGGTGGGAGTGGGAGGAGCGGGTGAATATGATCGGAACGATATCGAAGAGGAGGAGTA 120
DB 89 CTGGTGGGAGTGGGAGGAGGAGGAGTGAATATGATCGGAACGATATGCTTGAATGGGAGAT 148
QY 121 ACGCATCGAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGAGGAGATCGAAA 180
DB 149 GCCCGCGCAAAAGCCGGATTTAATACCGCATATGATCTTAAGAGAGAAAGGAGGAGATCG 208
QY 181 GACCTTGGCGCTTTTGAAGCGCGCATGTCTGATTAAGTATGTTGGTGGGTTAAAGGCTTAC 240
DB 209 GGCCTCGCGCTATAGG-GTTGGCGATGGCTGATTAAGTATGTTGGTGGGTTAAAGGCTTAC 267
QY 241 CAAGCGAGCGATCATAGTATGTTGGTGAAGAGAGCGACGACGCTGGGAGCTGAAGACG 300
DB 268 CAAGCGAGCGATCATAGTATGTTGGTGAAGAGAGCGACGACGCTGGGAGCTGAAGACG 327
QY 301 GCCCAGACTCTCTACCGGAGGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCCTGATC 360
DB 328 GCCCAGACTCTCTACCGGAGGAGCAGAGTGGGGAATTTTGGACAATGGGCGCAAGCCTGATC 387
QY 361 CAGCAATGCCGCGTGAAGTGAAGAGGCGCTTGGGTTTGAAGCTCTTTTCAATCGAAGACA 420
DB 388 CAGCAATGCCGCGTGAAGTGAAGAGGCGCTTGGGTTTGAAGCTCTTTTCAATCGAAGACA 447
QY 421 AAAAGTTACGGTAATTCGTGACTCAGTGAAGGATGCAAGAGAGAGAGAGAGCGGCTTAC 480
DB 448 AATCCCTGGCTCTAATATACGTGCGGAGGATGACGATACCGAAGAGATATGACACCGGCTTAC 507
QY 481 TACGTGCCAGCAGCGCGGCTAATACGTAGGAGTGAAGCCTTATCGAATTTACTGCGGCT 540
DB 508 TACGTGCCAGCAGCGCGGCTAATACGTAGGAGTGAAGCCTTATCGAATTTACTGCGGCT 567
QY 541 AAAAGGTGGCAGCGCGCTTTTGAAGTCAAGATGTGAATCCCGGCGCTTAACTTGGGAAT 600
DB 568 AAAAGGTGGCAGCGCGCTTTTGAAGTCAAGATGTGAATCCCGGCGCTTAACTTGGGAAT 627
QY 601 TGGCTTTGAATCTAAGAGCTGAGTGTGGCAGAGAGAGTGGAAATTCATGTGAAGAG 660
DB 628 TGGCTTTGAATCTAAGAGCTGAGTGTGGCAGAGAGAGTGGAAATTCATGTGAAGAG 687
QY 661 TGAATATGCTAGAGATATGAGAGAAACATCGATGCGAAGGAGCGCTCTGCGGTTAACT 720
DB 688 TGAATATGCTAGAGATATGAGAGAAACATCGATGCGAAGGAGCGCTCTGCGGTTAACT 747
QY 721 GACGCTCATGACGAAAGCGTGGGAGAGCAAGATTAAGATACCTGGTAGTCCACGCGC 780
DB 748 GACGCTCATGACGAAAGCGTGGGAGAGCAAGATTAAGATACCTGGTAGTCCACGCGC 807
QY 781 CTAAACGATGTCACTAGTGTGGGCTTATTAAGCTTGGTGAAGAGAGTAAACGCGGTA 840
DB 808 CTAAACGATGTCACTAGTGTGGGCTTATTAAGCTTGGTGAAGAGAGTAAACGCGGTA 867
QY 841 AATTGACCGCTGGGAGATACGCTGCGAAGATTAATACTCAAGAGAAATGACGGGAGACC 900
DB 868 AATTGACCGCTGGGAGATACGCTGCGAAGATTAATACTCAAGAGAAATGACGGGAGACC 927
QY 901 GCACAGCGGATGATTAATGATTAATTCATGCAACGCGAAGAACTTACCTTACCTT 960
DB 928 GCACAGCGGATGATTAATGATTAATTCATGCAACGCGAAGAACTTACCTTACCTT 987
QY 961 GACATGTAGCGAATTTTCTAGAGATTAATTAAGTCT--TCGGGAACGCTTACACAGG 1017
DB 988 GACATGTAGCGAATTTTCTAGAGATTAATTAAGTCT--TCGGGAACGCTTACACAGG 1047
QY 1018 CTGCATGGGCTGTGTCAGAGCTGTGTCGTAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1077
DB 1048 CTGCATGGGCTGTGTCAGAGCTGTGTCGTAGATGTTGGGTTAAGTCCCGCAACGAGCGCA 1107
QY 1078 ACCCTTGTATTATTGCAATCATTTTGTGGGCACTTTAATGAGACTCGCGGTGACAAA 1137

```

```

DB 1108 ACCCTTGTCTTACTTGTCTAC-----GCAGAGACTCTAAGAGAGACTGCCGCTGACAAA 1162
QY 1138 CCGAGAGAGAGTGGGAGATGACGTCAAGTCTCATGCGCTTATGGTATGGGCTTACACAG 1197
DB 1163 CCGAGAGAGAGTGGGAGATGACGTCAAGTCTCATGCGCTTATGGTATGGGCTTACACAG 1222
QY 1198 TAAATCAATGGCGGTACAGAGGTTGCCAACCCCGAGAGGAGGCTTAATCTCAAGAAC 1257
DB 1223 TCATACATATGTCGGAACAGAGGTTGCAACCCGCAAGGAGGAGCTTAATCTCAAGAAC 1282
QY 1258 GCGTGTAGTCCGAGTGGAGTCTGCAACTGCACTCGGTGAAGTCGAGATCTTAAT 1317
DB 1283 CCATGTATGTCGAGATTTGCACTTCTGCAACTGCACTGCACTGAATCGGAATGCTAAT 1342
QY 1318 CCGCATGACGATGTCCGCGTGAATACGTTCCGCGGCTTTTACACACGCGCGCTCACAC 1377
DB 1343 CCGCATGACGATGTCCGCGTGAATACGTTCCGCGGCTTTTACACACGCGCGCTCACAC 1402
QY 1378 CATGGAGTGGGTTTACACAGAGAGTATCTTAACCTTAAGAGAGGCGCTTGCACAG 1437
DB 1403 CATGGAGTGGGTTTACACAGAGAGTATCTTAACCTTAAGAGAGAGGCGCTTGCACAG 1462
QY 1438 TGAGATTGATGACTGGGCTG 1457
DB 1463 TAGATTGATGACTGGGCTG 1482

```

RESULT 32  
US-10-875-161-1

```

; Sequence 1, Application US/10875161
; Publication No. US2005009151A1
; GENERAL INFORMATION:
; APPLICANT: Chase, Matthew
; APPLICANT: Clayton, Robert
; APPLICANT: Landis, Bryan
; APPLICANT: Banerjee, Amit
; TITLE OF INVENTION: Methods for the Stereoselective Synthesis and Enantiomeric
; TITLE OF INVENTION: Enrichment of Beta-Amino Acids
; FILE REFERENCE: S0-3262-2-PR-US
; CURRENT APPLICATION NUMBER: US/10/875,161
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1
; LENGTH: 1530
; TYPE: DNA
; ORGANISM: Variororax paradoxus
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(1530)
; OTHER INFORMATION: 16S rRNA gene (rDNA) from Variororax paradoxus
; NAME/KEY: misc difference
; LOCATION: (1)..(1530)
; OTHER INFORMATION: alignment 0.75% different from 16S RNA gene (rDNA) of Variororax
; US-10-875-161-1

```

Query Match 78.1%; Score 1137.6; DB 8; Length 1530;

Best Local Similarity 87.5%; Pred. No. 4.5e-293; Indels 6; Gaps 3;

Matches 1278; Conservative 1; Mismatches 175; Indels 6; Gaps 3;

```

QY 1 ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGCGACACGAGATGCTTGCAT 60
DB 24 ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAACGCGACACGAGATGCTTGCAT 81
QY 61 CTGGTGGGAGTGGGAGGAGCGGGTGAATATGATCGGAACGATATCGAAGAGGAGGAGTA 120
DB 82 CTGGGCGGAGTGGGAGGAGGAGTGAATATGATCGGAACGATATCGGAGGAGGAGTA 141
QY 121 ACGCATCGAAGATGTCTAATACCGCATATTAAGTGAAGAGAGGAGGAGATCGAAA 180

```

142 ACGGAGGAAAGCTGTGCTAAATACCGCATACGATCTACGATGAAAGCAGGGGATCGAA 201  
QY 181 GACCTTGGCTTTTGGAGCGGCCCATGTCTGATTAGCTAGTTGGTGGGTTAAAGCCCTAC 240  
Db 202 GACCTTGGCCAAATGGAGCGGCCGATGTGCAATTAAGTGTGGTAAAGGCTTAC 261  
QY 241 CAAGCGCAGATCACTGATTGTTCTGAGAGGACGACGACCACTAGGGAAGCTGAGACAG 300  
Db 262 CAAGCTTTCGATCTGTAGCTGCTGAGAGGACGACGACCACTAGGGAAGCTGAGACAG 321  
QY 301 GCCAGACTCTTACCGGAGGAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCTGATC 360  
Db 322 GCCAGACTCTTACCGGAGGAGCAGTGGGAAATTTTGAACAATGGGCGCAAGCTGATC 381  
QY 361 CAGCAATGCCCGTGAAGGAAAGGCTTGGGTTGTTAAAGCTTTTCACTGAGAGAA 420  
Db 382 CAGCAATGCCCGTGAAGGAAAGGCTTGGGTTGTTAAAGCTTTTCACTGAGAGAA 441  
QY 421 AAAGTTACGGTAAATATCGTACTCATGACGGTATGACAGAAAGACACCGGCTAAC 480  
Db 442 AAAGCTTTTCTTAATTAAGAGGCTTAATGACGTAACCTTAAGAAATTAAGACCGGCTAAC 501  
QY 481 TACGTGCCAGACCGCGGTAATACGTAGGGTGCAGGCTTAATCGAATTAATCTGGGCGT 540  
Db 502 TACGTGCCAGACCGCGGTAATACGTAGGGTGCAGGCTTAATCGAATTAATCTGGGCGT 561  
QY 541 AAAGGTGCGCAGGGGCTTTGTAAGTGAATGTGAATCCCGGGCTTAACCTGGGAAAT 600  
Db 562 AAAGGTGCGCAGGGGCTTTGTAAGTGAATGTGAATCCCGGGCTTAACCTGGGAAAT 621  
QY 601 TGCGTTTAAACTCAAGGCTAGAGTGTGCGAGAGGAGTGAATTCATGTAGAGAG 660  
Db 622 TGCACTGTGACTGATAGCTAGAGTACGCTAGAGAGGAGTGAATTCATGTAGAGAG 681  
QY 661 TGAATGCGTGAAGTATGAAAGAAATCATGATGCGGAAGGACGCTCTGAGTAACT 720  
Db 682 TGAATGCGTGAAGTATGCGGAGAAACCGATGGGGAAGCAATCCCTGGAACCTGTACT 741  
QY 721 GACGCTCATGACCAAGACGTTGGGAGCAAAAGATTAAGTACCTTGGTGTGCTCAAGCC 780  
Db 742 GACGCTCATGACCAAGACGTTGGGAGCAAAAGATTAAGTACCTTGGTGTGCTCAAGCC 801  
QY 781 CTAAACGATCACTAGTATGTTGGGCTTATTAGGCTTGTGAGCAAGCTAACGCGTGA 840  
Db 802 CTAAACGATCACTAGTATGTTGGGCTTATTAGGCTTGTGAGCAAGCTAACGCGTGA 861  
QY 841 AGTTGACCGCTGGGAGTACGCTGCAAGATTAAGTCAAAAGGATTAAGCGGAGACC 900  
Db 862 AGTTGACCGCTGGGAGTACGCGCGCAAGGTTGAACTCAAAAGGATTAAGCGGAGACC 921  
QY 901 GCACAGCGGTGATTAATGTGATTAATTCATGCAAGCGGAAACCTTAACCTTACCTT 960  
Db 922 GCACAGCGGTGATTAATGTGATTAATTCATGCAAGCGGAAACCTTAACCTTACCTT 981  
QY 961 GACATGTAGCAATTTTCTAGAGATAGATAGTGTCT--TCGGGAAACGCTTAACAGGTG 1017  
Db 982 GACATGTAGCAATTTTCTAGAGATAGATAGTGTCTCAAAAGAGAACCTTAACAGGTG 1041  
QY 1018 CTGATGAGCTGTCTGACGCTGTGTCTGAGAGATGTTGGGTTAAAGTCCGCAACGAGCGCA 1077  
Db 1042 CTGATGAGCTGTCTGACGCTGTGTCTGAGAGATGTTGGGTTAAAGTCCGCAACGAGCGCA 1101  
QY 1078 ACCCTTGTCAATTAATGTCATCACTTTGGGCACTTTAATGAGCTGCGGAGCAAA 1137  
Db 1102 ACCCTTGTCAATTAATGTCATCACTTTGGGCACTTTAATGAGCTGCGGAGCAAA 1160  
QY 1138 CCGGAGGAAAGGTGGGAGTATGATCAAGTCTCATGAGCCCTTATGGGTTAGGCTTACAG 1197  
Db 1161 CCGGAGGAAAGGTGGGAGTATGATCAAGTCTCATGAGCCCTTATGGGTTAGGCTTACAG 1220  
QY 1198 TAATATCAATGCGGTATCAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1257  
Db 1221 TCATATCAATGCGGTATCAGAGGTTGCAACCCGCGAGGGGAGCTAATCTCAGAAAGC 1280

QY 1258 GCGTGTAGTCCGGATCGAGTCTGCAACTGCACTCCGTGAAGTCCGAAATGCTAGTAAT 1317  
Db 1281 CAGTGTAGTCCGGATCGAGTCTGCAACTGCACTCCGTGAAGTCCGAAATGCTAGTAAT 1340  
QY 1318 GCGGATCAGCATGTCCGCGTGAATTAACGTTCCCGGCTTTGTACACACCGCCGCTCAC 1377  
Db 1341 CGTGATCAGAAATGTCAACGAGTAATCGTTCCCGGCTTTGTACACACCGCCGCTCAC 1400  
QY 1378 CATGGAGTGGGTTTCAACAGAACAGTACTTAACCGTAAGAGAGGCGCTTCCACCG 1437  
Db 1401 CATGGAGCGGTTCTCCAGAAAGTATGCTTAACCGCAAGAGAGGCGGATTAACACCG 1460  
QY 1438 TGAGATTCACTAGTGGGTTG 1457  
Db 1461 CAGGTTCTGACTGGGTTG 1480

RESULT 33  
US-10-029-397A-39  
; Sequence 39, Application US/10029397A  
; Publication NO. US20030175709A1  
; GENERAL INFORMATION:  
; APPLICANT: MURPHY, GEORGE L.  
; APPLICANT: WHITLEY, J. PENN  
; TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING RNA POPULATIONS  
; FILE REFERENCE: AMBI.076US  
; CURRENT APPLICATION NUMBER: US/10/029,397A  
; NUMBER OF FILING DATE: 2002-03-19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 1485  
; TYPE: DNA  
; ORGANISM: Bordetella parapertussis  
US-10-029-397A-39

Query Match 78.1%; Score 1137.2; DB 6; Length 1485;  
Best Local Similarity 88.0%; Pred. No. 5,76-293;  
Matches 1287; Conservative 0; Mismatches 163; Indels 12; Gaps 4;

QY 1 ATTAAGCGCTGGCGGCAATGCTTTACACATGCAAGTGCAGCGGACGAGATGCTTCAT 60  
Db 1 ATTAAGCGCTGGCGGCAATGCTTTACACATGCAAGTGCAGCGGACGAGATGCTTCAT 58  
QY 61 CTGGTGGCAGTGGCGGAGGCTGTAATGCAATCGGAAGCTATCCAGAAAGGCGGTA 120  
Db 59 CTGGTGGCAGTGGCGGAGGCTGTAATGCAATCGGAAGCTATCCAGAAAGGCGGTA 118  
QY 121 ACGCATGAAGATGTCTAATAACGCAATTAATCTTAAGAGAGAAACAGAGGAGATCGAAA 180  
Db 119 ACTACGGAAAGCGTGGCTAATACCGCAATAGCCCTTACGGGGAGAAACGGGGGACTTTCC 178  
QY 181 GACCTTGGCTTTTGAAGCGCGCGGATGTGATTAAGTATGTTGGTGGGTTAAAGCTTAC 240  
Db 179 GACCTTGGCTTTTGAAGCGCGCGGATGTGATTAAGTATGTTGGTGGGTTAAAGCTTAC 238  
QY 241 CAAGCGCAGATCACTGATTGTTCTGAGAGAGCAACAGCAACCTGGGACTGAGACAG 300  
Db 239 CAAGCGCAGATCACTGATTGTTCTGAGAGAGCAACAGCAACCTGGGACTGAGACAG 298  
QY 301 GCCAGACTCTTACCGGAGGAGCAGAGTGGGAAATTTTGAACAATGGCGGCAAGCTGATC 360  
Db 299 GCCAGACTCTTACCGGAGGAGCAGAGTGGGAAATTTTGAACAATGGCGGCAAGCTGATC 358  
QY 361 CAGCAATGCCCGTGAAGGAAAGGCTTGGGTTGTTAAAGCTTTTCACTGAGAGAA 420  
Db 359 CAGCAATGCCCGTGAAGGAAAGGCTTGGGTTGTTAAAGCTTTTCACTGAGAGAA 418  
QY 421 AAAGTTACGGTAAATATGCTGACTCATGACGCTATGACAGAAAGAACCGGCTTAAC 480  
Db 419 AAAGTTACGGTAAATATGCTGACTCATGACGCTATGACAGAAAGAACCGGCTTAAC 478

QY	481	TACGTGCAGAGCGCGGTAAATGCTAGGGGTGCAAGGCTTAATCCGAAATTACAGGCGT	540
Db	479	TACGTGCAGAGCGCGGTAAATGCTAGGGGTGCAAGGCTTAATCCGAAATTACAGGCGT	538
QY	541	AAAGGTCGCGAGCGGCTTTGTAAAGTCAGATGTGAATCCCGGGCTTAACTTGGAAAT	600
Db	539	AAAGCGTCGCGAGCGGCTTTGGAAAGAAAGATGTGAATCCCGGGCTTAACTTTGGAAC	598
QY	601	TGCGTTGAAACTTACAGGCTTAGAGTGTGGCAGAGGAGGTGGAAATTCATGTGTACAG	660
Db	599	TGCAATTTTAACTACCGGGCTTAGAGTGTGTACAGAGGAGGTGGAAATTCGCGGTGTACAG	658
QY	661	TGAAATGCTAGAGTATGAAAGAACATCGATGGCGAAAGGAGCGCTCTGGGTTAACT	720
Db	659	TGAAATGCTAGAGTATGCGAGAGAACACCGATGGCGAGAGGAGCGCTCTGGGATTAACCT	718
QY	721	GACGCTCATGACAGAAACGTCGGGAGCAAAACAGATTTAGATACCTGTAGTACAGCC	780
Db	719	GACCTCATGACAGAAACGTCGGGAGCAAAACAGATTTAGATACCTGTAGTACAGCC	778
QY	781	CTAAACGATGTCAACTAGTGTGGGCTTATTAAGCTTGGTAAACGAAGCTAACGCGTGA	840
Db	779	CTAAACGATGTCAACTAGCTGTGGGCGCTTACGGGCTTGGTAGCGCAGCTAACGCGTGA	838
QY	841	AGTTGACCGCTCGGGAGATACGTCGCGAAGATTAAATCAAAGAAATTGACGGGGAGCC	900
Db	839	AGTTGACCGCTCGGGAGATACGTCGCGAAGATTAAATCAAAGAAATTGACGGGGAGCC	898
QY	901	GCACAAGCGGTGATATTATGTGGAATTAAATTGATGCAACGCGAAAAAAGCTTACTACCTT	960
Db	899	GCACAAGCGGTGATATTATGTGGAATTAAATTGATGCAACGCGAAAAAAGCTTACTACCTT	958
QY	961	GACATGTAGCGAAATTTCTAAGATAGATTAGTGTCT--TCGGAGAGCTTAACACAGGTG	1017
Db	959	GACATGTCTGAATCTCCGAAGAGATTTGGGAGTGTCTCGAAGAGAACCGGAACAGGTG	1018
QY	1018	CTGCAATGGCTCTCCGACGCTGTGTCGCGAATGTTGGGTTAAAGTCCCGAACAGAGCGCA	1077
Db	1019	CTGCAATGGCTCTCCGACGCTGTGTCGCGAATGTTGGGTTAAAGTCCCGAACAGAGCGCA	1078
QY	1078	ACCGTTGCAATTAATTTGGCATCATTTGGTTGGGCACTTAATGAGACTGCCGTGACAAA	1137
Db	1079	ACCGTTGCAATTAATTTGGCATCATTTGGTTGGGCACTTAATGAGACTGCCGTGTTACAA	1133
QY	1138	CCGAGAGAAAGTGGGATGACGTCAAGTCTTCATGAGCCCTTAATGGGTAGGGCTTACACG	1197
Db	1134	CCGAGAGAAAGTGGGATGACGTCAAGTCTTCATGAGCCCTTAATGGGTAGGGCTTACACG	1193
QY	1198	TAAATCAATGCGCGCTACAGAGGGTTCGCCAACCCCGAGGGGAGACTTAATCTCAGAAAGC	1257
Db	1194	TCATTCAAATGTCGCGGACAGAGGGTTCGCCAACCCCGAGGGGAGACCAATCCAGAAACC	1253
QY	1258	GCGTGTAGTCCGGAATCGGAATCTGCAACTGGAATCGGTGAAAGTGGATGCGTGTAAAT	1317
Db	1254	GCAATGTAGTCCGGAATCGGAATCTGCAACTGGAATCGGTGAAAGTGGATGCGTGTAAAT	1313
QY	1318	CGCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGGCTTGTACACACCGCCGTCACAC	1377
Db	1314	CGCGGATCAGCATGTGCGCGGTGAATACGTTCCCGGGCTTGTGTACACACCGCCGTCACAC	1373
QY	1378	CATGGAGATGGGTTTACCAAGAGCAGGTAGTCTAACCGTAA--GGAAGGCGCTTGGCCAC	1435
Db	1374	CATGGAGATGGGTTTATACCAAGAGTGTAGTCTTAACCGCAAGGGGGGGGCGATTACAC	1433
QY	1436	GGTGAGATTCAATGACTGGGGTG 1457	
Db	1434	GGTAGGATTCAATGACTGGGGTG 1455	

```

1 // GENERAL INFORMATION:
2 // APPLICANT: MURPHY, GEORGE L.
3 // APPLICANT: WHITLEY, J. PENN
4 // TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS
5 // FILE REFERENCE: AMB1:0761S
6 // CURRENT APPLICATION NUMBER: US/10/029,397A
7 // CURRENT FILING DATE: 2002-03-19
8 // NUMBER OF SEQ ID NOS: 73
9 // SOFTWARE: PatentIn Ver. 2.1
10 // SEQ ID NO 40
11 // LENGTH: 1464
12 // TYPE: DNA
13 // ORGANISM: Bordetella pertussis
14 // FEATURE:
15 // NAME/KEY: modified base
16 // LOCATION: (87) - (1391)
17 // OTHER INFORMATION: N = A, C, G or T/U
18 // US-10-029-397A-40

```

Query Match	77.7%;	Score 1131.4;	DB 6;	Length 1464;
Best Local Similarity	87.8%;	Pred. No. 2e-291;		
Matches 1268;	Conservative	0;	Mismatches 167;	Indels 10;
				Gaps 32;

Qy	1	ATTGAACGCTGGCGGCATGCTTTTACATGTGAATCTCAACGGGACGACGATCTGCAT	60
Db	27	ATTGAACGCTGGCGGCATGCTTTTACATGTGAATCTCAACGGGACGACGACG--GCTTCGGC	84
Qy	61	CTGTGTGGCAGATGGCGGACGGGGTAGTAATGCATCGAACGTAATCCGAAGAGGGGGGTA	120
Qy	121	ACGCATCCAAAGATGTGCTAATPACCGCATATACCTCTAAGAGAGAAACACAGGGGATCGAA	180
Db	145	ACTACGCGAAACGTAGCTAATACCGCATACGCCCTACGGGGGAAACGGGGGACCTTCG	204
Qy	181	GACCTTGCAGCTTTTGGACGGCGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGGCTTAC	240
Db	205	GGGCTTGCACTATTGGAGGGGCCGATATCGGATTAGCTGTTGGTGGGGTAACGGCTTAC	264
Qy	241	CAAGCGACGATCACTAGTAGTTGGTCTGAGAGGACGACCAACGCAACTGTGGGACTGACACG	300
Db	265	CAAGGAGACGATCCCTACTGGTTGTGAGAGGACACACGCACTGGGACTGAGACAG	324
Qy	301	GCCACAGACTCTACAGGGGAGGCGACGAGTGGGGGATTTTGGACAATGGGCGCACAGCTGATC	360
Db	325	GCCCAAGCTCTTACGGGAGGCGACGAGTGGGGATTTTGGACAATGGGGGCAACCTGATC	384
Qy	361	CAGCATGCGCGGTGAGTGAAGAAAGGCTTCGGGTTGTAAGCTCTTTCAGTGCAGAGA	420
Db	385	CAGCATCCCGGTGTGGAGTGAAGGCTTCGGGTTGTAAGCACTTTTGGCAGGAAGA	444
Qy	421	AAAGGTTACGTTAAATTAATCTGTACTCATGACGTTATCGACAGAAAGACCGGCGTAAC	480
Db	445	AACGGGACGGGTTAATATCTGTGTGCATCTGACGGTACTGTGAGAAATAGCACCGGCTTAC	504
Qy	481	TACGTCCAGACCGCGGGTAAATACGTAGGGTGCAGCGTTAATCGGAATTACTGCGGCGT	540
Db	505	TACGTCCAGACCGCGGGTAAATACGTAGGGTGCAGGGTTAATCGGAATTACTGCGGCGT	564
Qy	541	AAAGGTTGCGCAGCGCGCTTTGTAACTGCAATGTGAATCCCGGGCTTAACCTGGGAAT	600
Db	565	AAAGCGTGGCGCAGCGGTTTCGGAAGAAAGATGTGAATCCAGAGGGCTTTAACCTTGGAAC	624
Qy	601	TGGGTTGAAACCTACAAGGCTAGATGTGGACAGAGGAGGTGGAATTCATATGTATACAG	660
Db	625	TGATTTTAACTACCGGGCTAGATGTGTGTCAGAGGGAGGTGGAATTCCTCGGTGTACAG	684
Qy	661	TGAATATGCTAGAGATATGGAAGAACATCGATGGCGAAGGACGCTCTGSGGTTAACACT	720
Db	685	TGAATATGCTAGATATGCGAAGAACCATCGATGGCGAAGGACGCTCTGSGGTTAACACT	744
Qy	721	GACGCTCATGCAGAAAGCTGGGGAGCAACAGGATTAGTATACCTGTGATGTCAGCGC	780

Db 745 GACGCTCATGACGAAAGTGTGGGAGCAACAGATTAGATACCTGGTAGTCCACGCC 804  
Qy 761 CTAACGATGTCTAATTAAGTTGGGCTTATTAAGCTTGGTAAGAACTAACCGCTGA 840  
Db 805 CTAACGATGTCTAATTAAGTTGGGCTTATTAAGCTTGGTAAGAACTAACCGCTGA 864  
Qy 841 AGTTGACCGCTGGGAGATACCGTGCAGAAATTAACCAAGAAATTAAGCGGAGACC 900  
Db 865 AGTTGACCGCTGGGAGATACCGTGCAGAAATTAACCAAGAAATTAAGCGGAGACC 924  
Qy 901 GCACAGCGGTGATTAATGTGATTAATTCATGCAACGCAAAACCTTACCTTACCCTT 960  
Db 925 GCACAGCGGTGATTAATGTGATTAATTCATGCAACGCAAAACCTTACCTTACCCTT 984  
Qy 961 GACATGTACCGAATTTTCTAGATAGATTAAGTCT--TCGGGAACGCTAACACAGGTG 1017  
Db 985 GACATGTCTGGAATCCCGAAGAAATTTGGAGTGTCTGCGAAGAAACCGGAACAGAGTGTG 1044  
Qy 1018 CTGCAATGGCTGTCTGAGCTGTCTGTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCA 1077  
Db 1045 CTGCAATGGCTGTCTGAGCTGTCTGTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCA 1104  
Qy 1078 ACCCTTGTCTAATTAATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1137  
Db 1105 ACCCTTGTCTAATTAATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1159  
Qy 1138 CCGGAGAAAGTGGGAGATGACGTCAAGTCTCATGAGCCCTTATGCGTGAAGCTTACACG 1197  
Db 1160 CCGGAGAAAGTGGGAGATGACGTCAAGTCTCATGAGCCCTTATGCGTGAAGCTTACACG 1219  
Qy 1198 TAATCAATGAGCGCTGACAGAGGTTGACCAACCGGAGAGGAGAGCTATCTCAGAAAGC 1257  
Db 1220 TCATCAATGAGCGCTGACAGAGGTTGACCAACCGGAGAGGAGAGCTATCTCAGAAAGC 1279  
Qy 1258 GCGTGTGATGCGGATCGGAGTCTGCAACTGCACTCGTGAATGCGAATGCTAGTAT 1317  
Db 1280 GCGTGTGATGCGGATCGGAGTCTGCAACTGCACTCGTGAATGCGAATGCTAGTAT 1339  
Qy 1318 GCGGATGACATGTGCGGAGTGAATACGTTCCCGGCTTGTGACACACCGCCGCTACAC 1377  
Db 1340 GCGGATGACATGTGCGGAGTGAATACGTTCCCGGCTTGTGACACACCGCCGCTACAC 1399  
Qy 1378 CATGGAGTGGGTTTCAACGAAAGAGGATCTAAACGTTAGAGAGGCGCTGACAGG 1437  
Db 1400 CATGGAGTGGGTTTCAACGAAAGATGTTAGCTTAACGCAAGGAGGAGATTACACAGG 1459  
Qy 1438 TGAGA 1442  
Db 1460 TAGGA 1464

RESULT 35  
US-10-168-337A-3  
; Sequence 3, Application US/10168337A  
; Publication No. US20030170654A1  
; GENERAL INFORMATION:  
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholz, Philip; and Blackall,  
; APPLICANT: Linda L.  
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyphosphate  
; FILE REFERENCE: 002367PC/KF  
; CURRENT APPLICATION NUMBER: US/10/168,337A  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1478  
; TYPE: DNA  
; ORGANISM: Rhodocyclus purpureus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 78, 541, 761, 926, 1033, 1058, 1257  
; OTHER INFORMATION: n = unknown

US-10-168-337A-3  
Query Match 77.6%; Score 1130.6; DB 6; Length 1478;  
Best Local Similarity 87.4%; Pred. No. 3.3e-221;  
Matches 1269; Conservative 0; Mismatches 176; Indels 7; Gaps 3;  
Qy 1 ATTGACGCTGGCGCATGCTTTTACATGCAATGCAACCGGACGACGATGCTTGCAT 60  
Db 29 ATTGACGCTGGCGCATGCTTTTACATGCAATGCAACCGGACGATGCTTGCAT 86  
Qy 61 CTGTGGGAGTGGCGGACGCGGTGATATCATCGAAGCTATTCAGAAAGGAGGCTA 120  
Db 87 CGCGAAGCAAGTGGCGAAGCGGATGATATCATCGAAGCTATTCAGAAAGGAGGCTA 146  
Qy 121 ACGCATGGAAGATGTGTTATACCGCATATCTTAAAGGAGAAACGAGGGAGTTCGAA 180  
Db 147 ACGTACGGAAGATGTGTTATACCGCATATCTTAAAGGAGAAACGAGGGAGTTCGAA 206  
Qy 181 GACCTGGCTTTTGGAGCGGCGCATGTCTGATTAAGTGTGTTGGGTTAAAGGCTTAC 240  
Db 207 GACCTGGCTTTTGGAGCGGCGCATGTCTGATTAAGTGTGTTGGGTTAAAGGCTTAC 266  
Qy 241 CAAGGCGACATCAGTATGTTGTTCTGAGAGGACGACACGCACTGGGACTGAGACG 300  
Db 267 CAAGGCGACATCAGTATGTTGTTCTGAGAGGATGATCCGCACTGGGACTGAGACG 326  
Qy 301 GCCCAACTCTTACGCGGAGGACGACATGCGGAAATTTTGACAAATGGCGCAAGCTGATC 360  
Db 327 GCCCAACTCTTACGCGGAGGACGACATGCGGAAATTTTGACAAATGGCGCAAGCTGATC 386  
Qy 361 CAGCAATGCGCGTGAAGAGAGGCTTGGGTTTAAAGCTCTTTCAGTACAGAAAGA 420  
Db 387 CAGCAATGCGCGTGAAGAGAGGCTTGGGTTTAAAGCTCTTTCAGTACAGAAAGA 446  
Qy 421 AAAGGTTACGTTAAATATCTGTAATCTATGACGATGATGACAGAAAGACCGGCTTAC 480  
Db 447 AAAGGTTACGTTAAATATCTGTAATCTATGACGATGATGACAGAAAGACCGGCTTAC 506  
Qy 481 TACGTGCGACAGAGCGCGGTATACGTAAGGTGCAACCGTAAATTCGAGGCGT 540  
Db 507 TACGTGCGACAGAGCGCGGTATACGTAAGGTGCAACCGTAAATTCGAGGCGT 566  
Qy 541 AAAGGTTACGCGGCTTGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 567 AAAGGTTACGCGGCTTGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 626  
Qy 601 TGCGTTTGAACTCAAGAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTACAG 660  
Db 627 TGCGTTTGAGCTGCAAGAGCTAGAGTGTGCGAGAGGAGTGAATTCATGTGTACAG 686  
Qy 661 TGAAATGCGTGAATATGAAAGAACTCATGAGCGCAAGCGCTCTCGGTAAACACT 720  
Db 687 TGAAATGCGTGAATATGAAAGAACTCATGAGCGCAAGCGCTCTCGGTAAACACT 746  
Qy 721 GAGCTCATGACGAAAGCGGAGGAGCAACAGAGTTAGTACCTGTGATGACAGGC 780  
Db 747 GAGCTCATGACGAAAGCGGAGGAGCAACAGAGTTAGTACCTGTGATGACAGGC 806  
Qy 781 CTAACGATGTCTAATTAAGTTGGGCTTATTAAGCTTGGTAAGAACTAACCGCTGA 838  
Db 807 CTAACGATGTCTAATTAAGTTGGGCTTATTAAGCTTGGTAAGAACTAACCGCTGA 866  
Qy 839 GAAAGTTGACCGCTGGGAGATGACGTTGCAAGATTAACCAAGAAATTAAGCGGAGC 898  
Db 867 GAAAGTTGACCGCTGGGAGATGACGTTGCAAGATTAACCAAGAAATTAAGCGGAGC 926  
Qy 899 CCGCAAGGCGGTGATTAATGTGATTAATTCATGCAACGCAAAACCTTACCTACCC 958  
Db 927 CCGCAAGGCGGTGATTAATGTGATTAATTCATGCAACGCAAAACCTTACCTACCC 986  
Qy 959 TTGACATGTACGAAATTTTCTAGATAGATTAAGTGC---TTCCGGAACGCTAACAGAG 1015  
Db 987 TTGACATGTACGAAATTTTCTAGATAGATTAAGTGC---TTCCGGAACGCTAACAGAG 1046



[illegible]

QY	181	GACCTTCCGCTTTTGGAGCGCGCGAGTCTGATTAGCTAGTTGATGATGAGGAGGCTTAC	240
Db	177	GGCTTCCGCGAACCAGGCGCGCGAGTGGCAGATTAGTAGTGTGGTGGGCTTAAAGCTTAC	236
QY	241	CAAGGCACAGATCAGTAGTGTGCTGAGAGGAGCAGACCACACTTGGAGCTGAGACAGC	300
Db	237	CAACCTTACATCTGTAGCTGTGGTCTGAGAGGAGCAGACCAGCCACTTGGAGCTGAGACAGC	296
QY	301	GCCCAAGCTCTTACGGGAGGAGCAGCAGTGGGGAAATTTTGGACAAATGGGCGGCAACCTGATC	360
Db	297	GCCCAAGCTCTTACGGGAGGAGCAGCAGTGGGGAAATTTTGGACAAATGGGCGGCAACCTGATC	356
QY	361	CAGCAATGCCCGCGAGTAGAAGAAAGGCTTCGGGTTGTAAAGCTCTTTCAGTGAGAGAGA	420
Db	357	CAGCAATGCCCGCGAGTAGAAGAAAGGCTTCGGGTTGTAAAGCTCTTTCAGTGAGAGAGA	416
QY	421	AAAGGTTACGGTAAATTAATCGTGACTGACGCGTATCGACAGAGAGAGCAGCGGCTTAC	480
Db	417	AAAGGCTTCTTCTAATTAAGGGGGGTCATGACGGTACCGTAAAGATTAAGCAGCGGCTTAC	476
QY	481	TACGTGCCAGAGCGCCGGCTTAATTCGTAAGGTGCGAAGCGTTAATCGGAATTAACCTGGGCGT	540
Db	477	TACGTGCCAGAGCGCCGGCTTAATTCGTAAGGTGCGAAGCGTTAATCGGAATTAACCTGGGCGT	536
QY	541	AAAGGTCGCGAGCGCGCTTGTAAAGTCAGATGTGAATCCCGGCGCTTAACTGGGAAT	600
Db	537	AAAGGTCGCGAGCGCGCTTGTAAAGTCAGATGTGAATCCCGGCGCTTAACTGGGAAT	596
QY	601	TGCGTTTGAAATCTAACAGGCTAAGTGTGCGACAGGAGGTGGAAATTCATGTGTACAG	660
Db	597	TGCAATTTGTGCTGCATGCTGCGAGTACCGGACAGGGGGAGTGAATTCGCGGTGTACAG	656
QY	661	TGAATGTGCTAGAGATATGAAAGAACATCGATGCGAGAGGACGCTCTGGGTTAACT	720
Db	657	TGAATGTGCTAGATATGCGAGAGAACACCGATGCGAAGGCAATCCCTGGGCTGTAACT	716
QY	721	GACGCTCAGTGCAGGAAAGCGTGGGAGGCAAGAGATTAGTACCTTGTGTGTCAAGCC	780
Db	717	GACGCTCAGTGCAGGAAAGCGTGGGAGGCAAGAGATTAGTACCTTGTGTGTGTCAAGCC	776
QY	781	CTAAACGATGTCACTAGTGTGGGCTTATTAGGCTTGTGTAAAGAGCTTAAGCGGCTGA	840
Db	777	CTAAACGATGTCACTAGTGTGGGCTTACTGATCACTAATACGAAAGCTTAAGCGGCTGA	836
QY	841	AGTTAACCGCTGGGAGTAGCGGTGCGACAGATTAAACTCAAGAAATTGACGGGAGCC	900
Db	837	AGTTAACCGCTGGGAGTAGCGGTGCGACAGATTAAACTCAAGAAATTGACGGGAGCC	896
QY	901	GCAACAGCGGTGAGTTATGTGGAATTAATTCGATCGAACGCGAATACTTACCTTACCCTT	960
Db	897	GCAACAGCGGTGAGTTATGTGGAATTAATTCGATCGAACGCGAATACTTACCTTACCCTT	956
QY	961	GACATGTACGAATTTTCTAGAGATAGATTAGTGT---TGGGGACGCTAACAGAGTG	1017
Db	957	GACATGTACGAATCTTTTAGAGATAGAGATGTCTCAAAAGAGACCTTAACCAAGTG	1016
QY	1018	CTGCATGCGTGTCTCAGCTGCTGTGCGTAGATGTTGGGTTAAGTCCCGCAACGAGCGCA	1077
Db	1017	CTGCATGCGTGTCTCAGCTGCTGTGCGTAGATGTTGGGTTAAGTCCCGCAACGAGCGCA	1076
QY	1078	ACCTTGTCAATTAATTCGCATCATTTTGGTGGGCACTTAAATGAGACTGCGCGGTGACAA	1137
Db	1077	ACCTTGTCAATTAATTCGCATCATTTTGGTGGGCACTTAAATGAGACTGCGCGGTGACAA	1133
QY	1138	CCGAGAGAGGTGGGAGTAGCGTCAAGTCTCATATGCGCTTAATGGGTATGGGCTTACACG	1197
Db	1132	CCGAGAGAGGTGGGAGTAGCGTCAAGTCTCATATGCGCTTAAATGGGTATGGGCTTACACG	1191
QY	1198	TAAATCAATGGCGCTACAGAGAGGTTGCCAACCGCGAGGGGAGCTAATCTCAAGAAAGC	1257
Db	1192	TAAATCAATGGCGCTACAGAGAGGTTGCCAACCGCGAGGGGAGCTAATCCATMAAGC	1251

OY	1258	GCGTCGTAGTCCGGATCGGAGTCTGCACACTCCGACTCCGTGAAGTCGGAAATCGTTAGTAAT	1317
Db	1252	CAGTCGTAGTCCGGATCGGAGTCTGCACACTCCGACTCCGTGAAGTCGGAAATCGTTAGTAAT	1311
OY	1318	CGCGGATCAGCAATGTCGCGGATGAAATCGATTCCCGGGTCTTGTAACAACGCCCGCTCACAC	1377
Db	1312	CGCGGATCAGCAATGTCGCGGATGAAATCGATTCCCGGGTCTTGTAACAACGCCCGCTCACAC	1371
OY	1378	CATGGAGTAGGGTTTTACCAGAAGCAGGTAGTCTAACCCTAAGAGAGGGCGCTTGCACGG	1437
Db	1372	CATGGAGTAGGGTTTTACCAGAAGGTAGTAGCTTAACCGTAAGAGAGGGCGCTTACACGG	1431
OY	1438	TGAGATTCAATGACTGGGGTGG	1457
Db	1432	CAGGGTTCGTGACTGGGGTGG	1451
 RESULT 37 US-11-112-257-1 ; Sequence 1, Application US//11112257 ; Publication No. US20050196805A1 ; GENERAL INFORMATION: ; APPLICANT: Beck, James J ; APPLICANT: Barnett, Charles J ; APPLICANT: Yarnall, Michele S ; APPLICANT: Zetouni, Lillian ; TITLE OF INVENTION: Diagnostics for the detection of Acidovorax avenae subsp. ; TITLE OF INVENTION: citrullini, casual agen of bacterial fruit blotch in melons ; FILE REFERENCE: 7000IUNSP ; CURRENT APPLICATION NUMBER: US//11/112, 257 ; CURRENT FILING DATE: 2005-04-22 ; PRIOR APPLICATION NUMBER: US//10/397, 551 ; PRIOR FILING DATE: 2003-03-05 ; PRIOR APPLICATION NUMBER: 60/367, 628 ; PRIOR FILING DATE: 2002-03-25 ; NUMBER OF SEQ ID NOS: 43 ; SOFTWARE: PatentIn version 3.2 ; SEQ ID NO 1 ; LENGTH: 1481 ; TYPE: DNA ; ORGANISM: Acidovorax avenae subsp. citrullini US-11-112-257-1			
 Query Match      77.5%; Score 1129.2; DB 10; Length 1481; Beet Local Similarity    87.3%; Pred. No. 7.7e-291; Matches 1275; Conservative    0; Mismatches 173; Indels    12; Gaps    3;			
OY	1	ATTGAACGCTGGCGGCGCATGCTTTACATGCAATGGAACGGCAGACGGAGTCTTGCAAT	60
Db	1	ATTGAACGCTGGCGGCGCATGCTTTACATGCAATGGAACGGTMAACGGATCTTTCGGA---	57
OY	61	CTGGGAGGCGAGTGGCGGACGGGGAGTAAATGCATCGGAACGTATCCAGAAAGAGGGGGATA	120
Db	58	-TGCTGACAGATGGCGAACGGGTGAATATACATCGAAACGTGCCGATCGTGGGGGATA	116
OY	121	ACGCATCGAAGATGTGCTAATACCGCATATACTCTAAGAGAGAAAACAAGGGATCGAAA	180
Db	117	ACGAGGCGAAGCTTTTGCTAATACCGCATATAGTATGATGAAAGACAGGGGACCGTAA	176
OY	181	GACCTTGGCGCTTTTGGAGCGGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGGCTTAC	240
Db	177	GGCCTTGGCGCGAAGCGAGCGGCCGATGGAGATTAAGTAGTGGTGGGGTAAAGGCTTAC	236
OY	241	CAAGGCGAGATCAGTAGTGGTCTGAGAGAGACACACACACACTGGGACTGAGACAG	300
Db	237	CAGGCTACGATCTGTAGCTGGTCTGAGAGAGACACACACACACTGGGACTGAGACAG	296
OY	301	GCCCAGACTCTTAACGGGAGGACAGACTGGGGAATTTTGGACAATGGGCGCAAAGCTGATC	360
Db	297	GCCCAGACTCTTAACGGGAGGACAGACTGGGGAATTTTGGACAATGGGCGCAAAGCTGATC	356
OY	361	CAGCAATGCCGCTGAGTAGAAGAGGCTTTCGGGTTGTAAAGCTTTTCAAGTCGAGAGA	420
Db			

Db	357	CAGCCATGCCCGCTGCAGAGATGAAAGCCCTTCCGGTTGTAACCTGTTTGTAGCGAACA	416
Qy	421	AAAGGTATCCGGTAAATAATTCGTGACTGACGTATGACAGAAAGAACCCGGCTAAC	480
Db	417	AAAGCCTTCTTCTAATAAAGGGGGGTGATGACGGTACCGTAAAGATAAGACCCGGCTAAC	476
Qy	481	TACGTCGACAGACCGCCGGTAAATCGTAGGGGTGCAGAGGTTAATGGAATTACGTGGGGGT	540
Db	477	TACGTCGACAGACCGCCGGTAAATCGTAGGGGTGCAGAGGTTAATGGAATTACGTGGGGGT	536
Qy	541	AAAGGGTCCGACGAGCGGCTTTGTAAAGTCAGATGTGAAATCCCGGGCTTAACTTGGGAAT	600
Db	537	AAAGGTCGCGACGCGGTGATGTAAAGACAGATGTGAATCCCGGGCTCAACTTGGGAAC	596
Qy	601	TGCGTTTGAAACTACAGAGCTTAGAGTGTGGCAGAGGAGGTGGAATTCCATGTGTAGCAG	660
Db	597	TGCATTTTGTACTGCATCGCTGAGTACGGCAGAGGGGGATGGAATTCGCGGTGTAGCAG	656
Qy	661	TGAAATGGGTGAGATATGGAAGAACATGCATGGCGAAGGGACGCTCCGGGTTAACACT	720
Db	657	TGAAATGGGTGAGATATGCGAGAGAACACCGATGGCGAAGGGACATCCCTGGGCTGTACT	716
Qy	721	GACGCTCATGACAGAAACGTGGGGAGCAACAGGATAGATACCTGGTAGTCCAGCC	780
Db	717	GACGCTCATGACAGAAACGTGGGGAGCAACAGGATAGATACCTGGTAGTCCAGCC	776
Qy	781	CTAAACGATGTCAACTAGTTGTGGGCGCTTATAGGCTTGTGTAAAGAGCTAACCGCTGA	840
Db	777	CTAAACGATGTCAACTGGTTGTGGGTCTTCACTGACTCAATTAACGAAGCTAACCGCTGA	836
Qy	841	AGTTGACCGGCTGGGGAGTACGGTTCGCAAGTTAAATCAAAGAAATTGACGGGAGACC	900
Db	837	AGTTGACCGGCTGGGGAGTACGGGTCGCAAGGTGAACTCAAAGAAATTGACGGGAGACC	896
Qy	901	GCACAAGCGGTGAGATTATGTGATTAATTTCGATGCACGCGAAAAAAGCTTACCTT	960
Db	897	GCACAAGCGGTGAGATGTGTGTTTAATTTCATGTGCACGCGAAAAAAGCTTACCTT	956
Qy	961	GACATGTAGCGAAATTTTCTAGAGATAGATTAGTCT---TCGGGAAGCCTAACACAGGTG	1012
Db	957	GACATGTAGCGAAATCTTTAGAGATAGAGAGTGTCCGAAAGAGAACCGTAAACAGGTG	1016
Qy	1018	CTGCATGGCTGTCTGCAGCTGTGTCTGAGAAATGTTGGGTTAAATGCCCGCAACGAGCCA	1072
Db	1017	CTGCATGGCTGTCTGCAGCTGTGTCTGAGAAATGTTGGGTTAAATGCCCGCAACGAGCCA	1072
Qy	1078	ACCCTTGCAATTAATTTGCGCATCTTTGGTGGGCACTTAAATGAGACTGCGGTGACAAA	1132
Db	1077	ACCCTTGCAATTAATTTGCTGACAAA----GGGCACTTAATGGGACTGCCGTGACAAA	1132
Qy	1138	CCGAGGAAGGTGGGGATGACGTCAAGTCTCATGGGCCCTTAATGGGTAGGGCTTACACG	1192
Db	1132	CCGAGGAAGGTGGGGATGACGTCAAGTCTCATGGGCCCTTAATGGGTAGGGCTTACACG	1192
Qy	1198	TAAATCAATGCGCGGTACAGAGGGTTGCCAACCCGACGGGGAGCTTAATCTCAGAAAGC	1252
Db	1192	TCATACATATGCTGTGTACAGAGGGTTGCCAACCCGACGGGGAGCTTAATCCCAATAAAGC	1252
Qy	1258	GCGTCGTATGTCGGAATGGAGTCTGCACATCGACTCCGTTGAAGTCCGATCGCTAGTAT	1312
Db	1252	CAGTCGTATGTCGGAATGCCAGTCTGCACATCGACTCCGTTGAAGTCCGATCGCTAGTAT	1312
Qy	1318	CGCGGATCAGCATGTCCGCGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCAAC	1372
Db	1312	CGCGGATCAGAAATGTCCGCGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCAAC	1372
Qy	1378	CATGGGAGTGGGTTTCAACAGACAGGTAGCTTAACCGTAAAGAGGGCGCTTGCACAGG	1432
Db	1372	CATGGGAGCGGGTTCTGCAGAAAGTAGTACCTTAACCGTAAAGAGGGCGCTTACACAGG	1432
Qy	1438	TGAGATTCAATCTGGGGTG 1457	
Db	1432	CAGGGTTCGTACTGGGGTG 1451	

RESULT 38  
US-10-029-397A-42  
; Sequence 42, Application US/10029397A  
; Publication No. US20030175709A1  
; GENERAL INFORMATION:  
; APPLICANT: MURPHY, GEORGE L.  
; APPLICANT: WHITLEY, J. PENN  
; TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS  
; FILE REFERENCE: AMBI:076US  
; CURRENT APPLICATION NUMBER: US/10/029,397A  
; CURRENT FILING DATE: 2002-03-19  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 42  
; LENGTH: 1488  
; TYPE: DNA  
; ORGANISM: Burkholderia mallei  
US-10-029-397A-42

Query Match 77.3%; Score 1126.8; DB 6; Length 1488;  
Best Local Similarity 87.2%; Pred. No. 3,4e-290;  
Matches 1273; Conservative 0; Mismatches 177; Indels 10; Gaps 3;

QY 1 ATTGAACGCTGGCGCATGCTTTACATGCAAGTCGAAACGCGACGACGGATGCTTGCAT 60  
DB 3 ATTGAACGCTGGCGCATGCTTTACATGCAAGTCGAAACGCGACGACGG--GCTTCGGG 60  
QY 61 CTGTGGCGCATGTCGCGACGGGTGAATGATCGAAAGTATCCAAAGAGCGGCGTGA 120  
DB 61 CTGTGGCGCATGTCGCGACGGGTGAATGATCGAAAGTATCCAAAGATGCTGTGGCGGATA 120  
QY 121 AGCGATCGAAAGTGTCTTAATCCGCTATCTCTTAAGAGGAAAGCGGGGATGGA 180  
DB 121 GCGCGCGCAAGCGGATTAATACCGCATACGATCTGAGGATAAAGCGGGGACCTTCG 180  
QY 181 GACCTTGGCGCTTTGGAGCGCGCATGCTGATTAAGTATGTTGGTGGGTAAAGGCTTAC 240  
DB 181 GCGCTGCGCTTAAGGTTGGCGCATGCTTACATGCAAGTCGAAACGCGACGAC--GCTTCGG 240  
QY 241 CAAGCGACGATCATGATGTTGCTGAGAGGACGACACGCACTGGGACTGAGACG 300  
DB 241 CAAGCGACGATCATGATGTTGCTGAGAGGACGACACGCACTGGGACTGAGACG 300  
QY 301 GCCCAAGCTCTTACCGGAGGACGACGAGGGAATTTTGGACAAATGGCGGCAAGCTGATC 360  
DB 301 GCCCAAGCTCTTACCGGAGGACGACGAGGGAATTTTGGACAAATGGCGGCAAGCTGATC 360  
QY 361 CAGCAATGCGCGGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGGAGAGA 420  
DB 361 CAGCAATGCGCGGAGTGAAGAGGCTTCGGGTTGTAAGCTCTTTCAGTGGAGAGA 420  
QY 421 AAAGTTACCGTAAATATCTGATCATGACGGTATCGACAGAAAGACCGGGCTAAC 480  
DB 421 AAATCATTTCTGCTAATACCGGAGTGTATGCGTACCGGAAAGTAAGACACCGGCTAAC 480  
QY 481 TAGCTGCGACGACCGCGGTAAATCTGAGGTTGCAACCGTTAATCGGAATTAATCGGGCGT 540  
DB 481 TAGCTGCGACGACCGCGGTAAATCTGAGGTTGCAACCGTTAATCGGAATTAATCGGGCGT 540  
QY 541 AAAGGGTGGCAGCGCGCTTGTAAAGTCAATGTGAATCCCGGGCTTAACCTGGGAAT 600  
DB 541 AAAGCGTGGCAGCGCGCTTGTAAAGTCAATGTGAATCCCGGGCTTAACCTGGGAAT 600  
QY 601 TGCCTTTGAACCTCAAGGCTAGAGTGTGCGAGAGGAGGTGAATTCATGTGTACAG 660  
DB 601 TGCATTTGTACTGCGACGCTAGAGTATGCGAGAGGAGGTGAATTCATGTGTACAG 660  
QY 661 TGAATGCGTGAATATGAAGAACATCGATGCGGAAGGACGCTCTGTGGTTAAACAT 720  
DB 661 TGAATGCGTGAATATGAAGAACATCGATGCGGAAGGACGCTCTGTGGTTAAACAT 720

QY 721 GACGCTCATGACCAAGAGCGTGGGAGCAACAGGATTAGATACCTGTAATCCAGCC 780  
DB 721 GACGCTCATGACCAAGAGCGTGGGAGCAACAGGATTAGATACCTGTAATCCAGCC 780  
QY 781 CTAAACGATGTCATTAATGTTGGGCTTTAATAGCTTGTGAACGAAGCTAACCGCTGA 840  
DB 781 CTAAACGATGTCATTAATGTTGGGCTTTAATAGCTTGTGAACGAAGCTAACCGCTGA 840  
QY 841 AGTTACCGCTGGGAGTACCGGTGCAAGATTAATTAACGAAAGGAAATGAGCGGAGACC 900  
DB 841 AGTTACCGCTGGGAGTACCGGTGCAAGATTAATTAACGAAAGGAAATGAGCGGAGACC 900  
QY 901 GCACAAAGCGTGGATTAATGATTAATTCATGCAACGCGAAACCTTACCTACCTT 960  
DB 901 GCACAAAGCGTGGATTAATGATTAATTCATGCAACGCGAAACCTTACCTACCTT 960  
QY 961 GACATGTAGCAATTTTCTAGATAGATTAATGCT---TCGGAAACCTTAACACAGGTG 1017  
DB 961 GACATGTAGCAAGCCCATAGATTAATGCTGCTCGAAAGAGAACCGGCGACAGGTG 1020  
QY 1018 CTGATAGGCTGTCTGACGCTGCTGTGAGATGTTGGTTAATGTCGCGCAACGAGCGCA 1077  
DB 1021 CTGATAGGCTGTCTGACGCTGCTGTGAGATGTTGGTTAATGTCGCGCAACGAGCGCA 1080  
QY 1078 ACCCTTGTCAATTAATGCTCAATGCTTGTGGGCACTTAATGAGACTCGGCTGACAA 1137  
DB 1081 ACCCTTGTCTTAATGCTCAATGCTTGTGGGCACTTAATGAGACTCGGCTGACAA 1135  
QY 1138 CCGGAGGAAGTGGGAGTACGCTCAAGTCTCATGAGCCTTAATGAGGCTTCAACAG 1197  
DB 1136 CCGGAGGAAGTGGGAGTACGCTCAAGTCTCATGAGCCTTAATGAGGCTTCAACAG 1195  
QY 1198 TAATCAATGAGCGCTGACAGAGGTTGCCAACCCCGAGCGGAGCTAATCTCAAGAAC 1257  
DB 1196 TAATCAATGAGCGCTGACAGAGGTTGCCAACCCCGAGCGGAGCTAATCTCAAGAAC 1255  
QY 1258 GCGTGTATCCGAGTCGAGATCTGCAATCGAATCCGGAAGTGGGAATGCTAGTAAT 1317  
DB 1256 GCGTGTATCCGAGTCGAGATCTGCAATCGAATCGAATGGAATGCTAGTAAT 1315  
QY 1318 CCGGATCAGCATGTGCGCGGTGAATGCTCCGGGCTTTGTAACACACCGCGGTACAC 1377  
DB 1316 CCGGATCAGCATGTGCGCGGTGAATGCTCCGGGCTTTGTAACACACCGCGGTACAC 1375  
QY 1378 CATGGAGTGGGTTTCAACCAAGAGGTAATCTTAACGTAAGAGAGGCGCTTGCACCG 1437  
DB 1376 CATGGAGTGGGTTTCAACCAAGAGGTAATCTTAACGTAAGAGAGGCGCTTGCACCG 1435  
QY 1438 TGAGATTCAATGACTGGGCTG 1457  
DB 1436 TAGGATTCAATGACTGGGCTG 1455

RESULT 39  
US-10-029-397A-45  
; Sequence 45, Application US/10029397A  
; Publication No. US20030175709A1  
; GENERAL INFORMATION:  
; APPLICANT: MURPHY, GEORGE L.  
; APPLICANT: WHITLEY, J. PENN  
; TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS  
; FILE REFERENCE: AMBI:076US  
; CURRENT APPLICATION NUMBER: US/10/029,397A  
; CURRENT FILING DATE: 2002-03-19  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 45  
; LENGTH: 1544  
; TYPE: DNA  
; ORGANISM: Neisseria meningitidis  
US-10-029-397A-45

Query Match 77.2%; Score 1124.6; DB 6; Length 1544;



QY 1 ATTGAACGCTGGCGGATGCTTTTACATGCAAGTCGACCGGACGAC--GGATGCTTGC 58  
DB 307743 ATTGAACGCTGGCGGATGCTTTTACATGCAAGTCGACCGGACGACGAGAAAGCTTGC 307802  
QY 59 ATCT--GGTGGGAGTGGCGGAGCGGAGTGAATGATGGAACGTTATCCAAAGAGGGG 116  
DB 307803 TTCTCGGGTGGCGAGTGGGAAACGGGTGAGTAAATATCGAAGTGAACGAGTATGAGGGG 307862  
QY 117 GGTAAACGATCGAANAATGCTTAATACCGCATATATCTTAAGGAGAAAGAGGGGATC 176  
DB 307863 GATTAATGATCGAANAATGCTTAATACCGCATATATCTTAAGGAGAAAGAGGGGATC 307922  
QY 177 GAAAGACTTGGCGCTTTTGAAGCGGCGGATGCTGATTAAGTATGTTGGGTAAAGGC 236  
DB 307923 TTCCGGGCTTGGCGCTTTTGAAGCGGCGGATGCTGATTAAGTATGTTGGGTAAAGGC 307982  
QY 237 CTACCAAGGCGGATGCTGATTAAGTATGTTGGGTAAAGGC 296  
DB 307983 CTACCAAGGCGGATGCTGATTAAGTATGTTGGGTAAAGGC 308042  
QY 297 CACGGCCGATCTCTTAACGGGAGGCGAGTGGGGAATTTTGGCAATGGGCGCAAGCT 356  
DB 308043 CACGGCCGATCTCTTAACGGGAGGCGAGTGGGGAATTTTGGCAATGGGCGCAAGCT 308102  
QY 357 GATCCAGCAATGCGCGGTGATGAAGAAAGGCTTCCGGTGTAAAGCTCTTTCAGTCAAG 416  
DB 308103 GATCCAGCAATGCGCGGTGATGAAGAAAGGCTTCCGGTGTAAAGCTCTTTCAGTCAAG 308162  
QY 417 AAGAAAGGTTACGTTAATATCTGATCTATGACGGTATCGACAGAAAGACCCGC 476  
DB 308163 AAGAAAGGTTACGTTAATATCTGATCTATGACGGTATCGACAGAAAGACCCGC 308222  
QY 477 TAACTACGTCGCGGAGCGGCGGTAATAGTAGGGTGAAGGCTTAATCGAATTAATCG 536  
DB 308223 TAACTACGTCGCGGAGCGGCGGTAATAGTAGGGTGAAGGCTTAATCGAATTAATCG 308282  
QY 537 GCGTAAAGGAGTGGCGAGCGGCTTTGTAAAGTCAATGTAATATCCCGGCTTAACCTCG 596  
DB 308283 GCGTAAAGGAGTGGCGAGCGGCTTTGTAAAGTCAATGTAATATCCCGGCTTAACCTCG 308342  
QY 597 GAATTCGCTTTGAACTACAAAGGCTTAAGTGTGGCGAGGAGGAGTGAATTCATGTGA 656  
DB 308343 GAATTCGCTTTGAACTACAAAGGCTTAAGTGTGGCGAGGAGGAGTGAATTCATGTGA 308402  
QY 657 GCAGTGAATGCTTAAGTGAAGTGAAGAAATATGATGCGAAGGCGCTTCCGGTGA 716  
DB 308403 GCAGTGAATGCTTAAGTGAAGTGAAGAAATATGATGCGAAGGCGCTTCCGGTGA 308462  
QY 717 CACTGACGCTCATGACGAAGGCGTGGGAGCAAAACAGATTAATCTCTGTAAGTCA 776  
DB 308463 CACTGACGCTCATGACGAAGGCGTGGGAGCAAAACAGATTAATCTCTGTAAGTCA 308522  
QY 777 CGCCCTAAACGATGTCATAGTGTGGCTTATTA--GGCTTGTGAAGAACTTAACG 835  
DB 308523 CGCCCTAAACGATGTCATAGTGTGGCTTATTA--GGCTTGTGAAGAACTTAACG 308582  
QY 836 CGTGAAGTGAACGCTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 895  
DB 308583 CGTGAAGTGAACGCTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 308642  
QY 896 GACCCGCAAGAGCGGTGATTAATGATTAATTCATGCAACGCGAAGAACTTACCTTA 955  
DB 308643 GACCCGCAAGAGCGGTGATTAATGATTAATTCATGCAACGCGAAGAACTTACCTTA 308702  
QY 956 CCCTTGAATGATGCAATTTTCTAAGATTAAGTATG--CTTGGGAAAGCTTAACAG 1014  
DB 308703 GTCTTGAATGATGCAATTTTCTAAGATTAAGTATG--CTTGGGAAAGCTTAACAG 308762  
QY 1015 GTCTTGAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1074  
DB 308763 GTCTTGAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 308822

QY 1075 GCAACCTTGTATTAATGTCATATGTTGGTGGGCACTTAAATGAGATGCGGATGAC 1134  
DB 308823 GCAACCTTGTATTAATGTCATATGTTGGTGGGCACTTAAATGAGATGCGGATGAC 308882  
QY 1135 AAACCGAGGAAGTGGGAGTGAAGTCAAGTCTTATGAGGCTTATGAGGCTTAC 1194  
DB 308883 AAACCGAGGAAGTGGGAGTGAAGTCAAGTCTTATGAGGCTTATGAGGCTTAC 308942  
QY 1195 AGTAAATCAATGCGGCTTACAGAGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAA 1254  
DB 308943 AGTAAATCAATGCGGCTTACAGAGGTTGCCAACCCGCGAGGGGAGCTTAATTCAGAA 309002  
QY 1255 AGCGGTGATGCTCGGATGCGAGTCTGCAATGCTGACCTCGGAGTGGAAATCTGAT 1314  
DB 309003 AACCGATGATGCTCGGATGCGAGTCTGCAATGCTGACCTCGGAGTGGAAATCTGAT 309062  
QY 1315 AATCGGATGAGATGCTGCGGATGAATGCTTCCGGGTCTTGTACACACCGCCGTCA 1374  
DB 309063 AATCGGATGAGATGCTGCGGATGAATGCTTCCGGGTCTTGTACACACCGCCGTCA 309122  
QY 1375 CACCATGGGAGTGGGTTTACCAAGAGGATGATCTTAACCTTAAGAGGCGCTTCCA 1434  
DB 309123 CACCATGGGAGTGGGTTTACCAAGAGGATGATCTTAACCTTAAGAGGCGCTTCCA 309182  
QY 1435 CGGTGATGATTCATGACTGGGCTG 1457  
DB 309183 CGGTGATGATTCATGACTGGGCTG 309205

## RESULT 41

US-10-915-740A-1068/C  
; Sequence 1068, Application US/10915740A  
; Publication No. US2005019136A1

## GENERAL INFORMATION:

; APPLICANT: Frazer, Claire M.  
; APPLICANT: Hickey, Erin  
; APPLICANT: Peterson, Jeremy  
; APPLICANT: Tettelein, Herve  
; APPLICANT: Venter, J. Craig  
; APPLICANT: Masignani, Vega  
; APPLICANT: Galeotti, Cesira  
; APPLICANT: Mora, Manroza  
; APPLICANT: Ratti, Giulio  
; APPLICANT: Scarselli, Maria  
; APPLICANT: Scariato, Vincenzo  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizzi, Mariagrazia  
; APPLICANT: Grandi, Guido  
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use  
; FILE REFERENCE: 002441.00090  
; CURRENT APPLICATION NUMBER: US/10/915,740A  
; PRIOR FILING DATE: 2004-08-11  
; PRIOR APPLICATION NUMBER: 09/806,866  
; PRIOR FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: USN 60/103,794  
; PRIOR FILING DATE: 1998-10-09  
; PRIOR APPLICATION NUMBER: USSN 60/132,068  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: PCT/US99/25373  
; PRIOR FILING DATE: 1999-10-08  
; NUMBER OF SEQ ID NOS: 1068  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1068  
; LENGTH: 2242716  
; TYPE: DNA  
; ORGANISM: Neisseria meningitidis  
; US-10-915-740A-1068

## Query Match

Best Local Similarity 87.4%; Score 1124.6; DB 9; Length 2242716;  
Matches 1278; Conservative 0; Mismatches 179; Indels 6; Gaps 4;

QY 1 ATTGAACGCTGGCGGATGCTTTTACATGCAAGTGAAGGAGGAC--GGATGCTTGC 58

```
Db 1700902 ATTTGACGCTGGCGGCGATCTTTTACACATGCAAGTCGACGGCAGACAGAAAGCTTGC 1700843
Qy 59 ATCT--GGTGGCGATGGCGGACGGGTGATGATGATGGAACCTATTCAGAAAGGGG 116
Db 1700842 TTCTGGGGTGGGAGATGGGGAACGGGTGATGATGATGGAACCTATTCAGATGATG 1700783
Qy 117 GGTAAAGCATGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 176
Db 1700782 GATTAACGTATGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1700723
Qy 177 GAAAGACCTTGGCGCTTTTGGAGCGGCGCATGCTGATGATGATGATGATGATGATG 236
Db 1700722 TTGGGGCGCTTGGCGCTTTTGGAGCGGCGCATGCTGATGATGATGATGATGATG 1700663
Qy 237 CTACCAAGGCGACGATCAGTATGATGATGATGATGATGATGATGATGATGATGATG 296
Db 1700662 CTACCAAGGCGACGATCAGTATGATGATGATGATGATGATGATGATGATGATG 1700603
Qy 297 CACGGCGCGACGATCCTTACCGGGAGGACGATGAGGGGATTTTGGACATGGGGCGAAG 356
Db 1700602 CACGGCGCGACGATCCTTACCGGGAGGACGATGAGGGGATTTTGGACATGGGGCGA 1700543
Qy 357 GATCCAGCATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 416
Db 1700542 GATCCAGCATGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 1700483
Qy 417 AAGAAAGGTTACGGTAAATTAATCTGATCATAAGCGGTATGACAGAAAGAACACCG 476
Db 1700482 AAGAAAGGTTACGGTAAATTAATCAACGGGTATGACAGGTATGACAGAAATTAAG 1700423
Qy 477 TAACTACGTGCGACGACGCGCGGTAAATGATGATGATGATGATGATGATGATGATG 536
Db 1700422 TAACTACGTGCGACGACGCGCGGTAAATGATGATGATGATGATGATGATGATG 1700363
Qy 537 GCGTAAAGGATGCGACGCGCGCTTTGATGATGATGATGATGATGATGATGATGATG 596
Db 1700362 GCGTAAAGGATGCGACGCGCGCTTTGATGATGATGATGATGATGATGATGATG 1700303
Qy 597 GAATTCGCTTTGAAATCTACAGCGCTAAGTGTGCGACAGGAGGTGAAATTCATGT 656
Db 1700302 GAATTCGCTTTGAAATCTGAGTGTGCTGAGTGTGCGAGGAGGTGAAATTC 1700243
Qy 657 GCAGTAAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 716
Db 1700242 GCAGTAAATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1700183
Qy 717 CACTGACGCTCATGACGAAAGCGTGGGAGCAACAGGATTAATGATGATGATGATG 776
Db 1700182 CACTGACGCTCATGACGAAAGCGTGGGAGCAACAGGATTAATGATGATGATG 1700123
Qy 777 CGCCCTAAACGATGCTAATGATGATGATGATGATGATGATGATGATGATGATG 835
Db 1700122 CGCCCTAAACGATGCTAATGATGATGATGATGATGATGATGATGATGATG 1700063
Qy 836 CGTGAAGTTGACCGGCTGGGAGTACCGTCCGAAATTAATTAATTAATTAATTA 895
Db 1700062 CGTGAAGTTGACCGGCTGGGAGTACCGTCCGAAATTAATTAATTAATTA 1700003
Qy 896 GACCGGCAAGCGGCTGATTAATGATGATGATGATGATGATGATGATGATGATG 955
Db 1700002 GACCGGCAAGCGGCTGATTAATGATGATGATGATGATGATGATGATGATG 1699943
Qy 956 CCTTGAATGATGAGGAAATTTTCTAGAGATGATGATGATGATGATGATGATGATG 1014
Db 1699942 GCTTGAATGATGAGGAAATTTTCTAGAGATGATGATGATGATGATGATGATG 1699883
Qy 1015 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1074
Db 1699882 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1699823
Qy 1075 GCAACCTTGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1134
```

```
Db 1699822 GCAACCTTGCATTAATGATGATGATGATGATGATGATGATGATGATGATG 1699763
Qy 1135 AACCGGAGGAAAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1194
Db 1699762 AAGCGGAGGAAAGTGGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1699703
Qy 1195 ACGTAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1254
Db 1699702 ACGTAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1699643
Qy 1255 AGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1314
Db 1699642 AAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1699583
Qy 1315 AATCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1374
Db 1699582 AATCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1699523
Qy 1375 CACCATGGAAGTGGGATTTTACAGAGAGGATGATGATGATGATGATGATGATG 1434
Db 1699522 CACCATGGAAGTGGGATTTTACAGAGAGGATGATGATGATGATGATGATGATG 1699463
Qy 1435 CGGTGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1497
Db 1699462 CGGTGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1699400

RESULT 42
US-10-029-397A-43
; Sequence 43, Application US/10029397A
; Publication No. US20030175709A1
; GENERAL INFORMATION:
; APPLICANT: MURPHY, GEORGE L.
; APPLICANT: WHITLEY, J. PENN
; TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS
; FILE REFERENCE: AMBI:076US
; CURRENT APPLICATION NUMBER: US/10/029,397A
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1610
; TYPE: DNA
; ORGANISM: Burkholderia pseudomallei
US-10-029-397A-43

Query Match 77.0%; Score 1122; DB 6; Length 1610;
Best Local Similarity 87.0%; Pred. No. 6; Se-289;
Matches 1270; Conservative 0; Mismatches 180; Indels 10; Gaps 3;

Qy 1 ATTTGAAGCTGGCGGATGCTTTTACACATGCAAGTCGACGGCAGACGATGCTTGCAT 60
Db 82 ATTTGAAGCTGGCGGATGCTTTTACACATGCAAGTCGACGGCAGACGATGCTTGCAT 139
Qy 61 CTGGTGGCGAGTGGCGGACGGGTGATGATGATGATGATGATGATGATGATGATGATG 120
Db 140 CTGGTGGCGAGTGGCGGACGGGTGATGATGATGATGATGATGATGATGATGATGATG 199
Qy 140 CTGGTGGCGAGTGGCGGACGGGTGATGATGATGATGATGATGATGATGATGATGATG 199
Db 121 ACGCATGAAAGATGCTAATACCGCATATCTTAAGAGAAAGCAGGGGATGCAAA 180
Qy 200 GCCCGGAAAGCCGAATTAATACCGCATAGATGATGATGATGATGATGATGATGATG 259
Db 181 GACCTTGGCTTTTGGAGCGGCGATGATGATGATGATGATGATGATGATGATGATGATG 240
Qy 260 GGCTTCGCGCTAATGAGGTGGCCGATGCTGATTAATGATGATGATGATGATGATG 319
Db 241 CAAGGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Qy 320 CAAGGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 379
Db 301 GCCCAGACTCTTACGGGAGGACGATGAGGAAATTTTGAATGATGATGATGATGATG 360
Qy 380 GCCCAGACTCTTACGGGAGGACGATGAGGAAATTTTGAATGATGATGATGATGATG 439
```

Oy	361	CAGCAATGCGCGGTGAAGAAAGGCGCTTGAGGTGTAAGACTTTTCAGTCGAGAAGA	420
Db	440	CAGCAATGCGCGGTGTGAAGAAAGGCGCTTGAGGTGTAAGACTTTTCGCGAAGA	439
Oy	421	AAAGGTTACGGTAAATATATCGTGACTCAAGAAGGATTCGACAGAAGAAGCAACGGCTAAC	480
Db	500	AATCATTTCTGGCTAATATACCCGAGTGAATGACGGGTACCGGAAGAATTAAGCAACCGGCTAAC	559
Oy	481	TACGTGCGACAGCGCCGCGGTAAATCTGATGGGTGCAAGCGTTAATCCGAATTACCGAGGCGT	540
Db	560	TACGTGCGACAGCGCGCGGTAAATCTGATGGGTGCAAGCGTTAATCCGGAATTACCGAGGCGT	619
Oy	541	AAAGGGTCCGACAGCGCGCTTTGTAAAGTCAGATGTGAATCCCGGCGCTTAACTTGGAAT	600
Db	620	AAAGGTCGCCAGAGCGGCTTTGCTAAGACCGATGTGAATCCCGGCGCTCAACTTGGAAT	679
Oy	601	TGCGTTTGAATCTACAGGCTTAAGATGTGGCAAGGAGGTGTGAATTTCAATGTGTACAG	660
Db	680	TGCAATGTGTGACTGGCAGGCTTAAGATGAGCAAGGCGGGGTGAAGAAATTCACAGTGTACAG	739
Oy	661	TGAATGTGGTGAAGATATGSAAGAACATCGATGGCGAAGGAGGACCTCCGSGTTAACT	720
Db	740	TGAATGTGGTGAAGATATGSAAGAACATCGATGGCGAAGGAGGACCTCCGSGCCAAATACT	799
Oy	721	GACGCTCAGACGAAACGCTGGGAGCAAAACAGATTAAGATACCTGTGTAGTCACGCC	780
Db	800	GACGCTCAGACGAAACGCTGGGAGCAAAACAGATTAAGATACCTGTGTAGTCACGCC	859
Oy	781	CTAAACGATGTCAACTAGTTGTGGGCGCTTATAGGCTTGGTTAACGAAGCTAACGCGTGA	840
Db	860	CTAAACGATGTCAACTAGTTGTGGGAGTTCAATTTCTTATATACGTAGCTTAACGCGCGGA	919
Oy	841	AGTTACACGCGCTGGGAGATACGCGTGCAGAAATTAAACTCAAGAGAAATTGAACGGGAGACC	900
Db	920	AGTTACACGCGCTGGGAGATACGCGTGCAGAAATTAAACTCAAGAGAAATTGAACGGGAGACC	979
Oy	901	GCAACAAGCGGTGATTAATGTGATTAATTCATGTGCAACGCGAAAACCTTACCTTACCTT	960
Db	980	GCAACAAGCGGTGATTAATGTGATTAATTCATGTGCAACGCGAAAACCTTACCTTACCTT	1039
Oy	961	GACATGTAGCGAATTTTTCAGAGATAGAAATTAGTGTCT--TCGGGAAACCTTAACAGAGTG	1017
Db	1040	GACATGTAGCGAAGCCCGATGAGAGTTGGGCGTGTCCAAAGAGAAACCGGCGCACAGTGT	1039
Oy	1018	CTGCAATGCGCTGTCTCAAGCTCGTGTCTGAGATGTTGGGTTAATGTCCGCGCAACGACGCA	1077
Db	1100	CTGCAATGCGCTGTCTCAAGCTCGTGTCTGAGATGTTGGGTTAATGTCCGCGCAACGACGCA	1159
Oy	1078	ACCCCTTGTCAATTAATTTGCCATCAATTTGGTGGGCACTTAAATGAACTGCCGTGACAA	1137
Db	1160	ACCCCTTGTCTTAATTTGGTGTAC-----GCMAAGGCACTTAAGGAACCTGCCGTGACAA	1214
Oy	1138	CCGAGAGAGATGGGGATGACGTCAAAGCTCTCATGAGCCCTTAATGGGTAGGGCTTACACG	1197
Db	1215	CCGAGAGAGATGGGGATGACGTCAAAGCTCTCATGAGCCCTTAATGGGTAGGGCTTACACG	1274
Oy	1198	TAATACATAGCGCGCTTACAGAGGGTTGCCAACCCCGAGGGGGAGCTAATCTCAGAAAC	1255
Db	1275	TAATACATAGCGCGCTTACAGAGGGTTGCCAACCCCGAGGGGGAGCTAATCTCAGAAAC	1334
Oy	1258	GCGTCGTAGCTCCGGAATCGAAGCTGGAACCTGCACTCCGTAAGTGGGAATCGCTAGTAAT	1317
Db	1335	GCGTCGTAGCTCCGGAATCGAAGCTGGAACCTGCACTCCGTAAGTGGGAATCGCTAGTAAT	1399
Oy	1318	CGCGGATTCAGCAATGTCCGCGTGAATAGTTCCCGGGTCTTGTACACACCGCCCGTACAC	1377
Db	1395	CGCGGATTCAGCAATGTCCGCGTGAATAGTTCCCGGGTCTTGTACACACCGCCCGTACAC	1454
Oy	1378	CATGGAGTGGGTTTACAGAAAGCAAGATAGTCTAACCGTAAGAGGGGGCGCTTGCACAG	1433
Db	1455	CATGGAGTGGGTTTACAGAAAGTGGCTTAACTCCGAAGAGAGAGCTTACACAG	1511

[illegible]



```
Qy 657 GCAGTGAATGCGTGAAGATATGGAAGAACATCGATGCGAAGGACGCTCTGGGTTAA 716
Db 689 GCAGTGAATGCGTGAAGATATGGAAGAACATCGATGCGAAGGACGCTCTGGGTTAA 748
Qy 717 CACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTTAGTACCCTGGTATGCA 776
Db 749 CACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTTAGTACCCTGGTATGCA 808
Qy 777 CGCCCTAAACGATGCACTAGTGTGGGCTTTATTA-GGCTTGTAACGAAGCTAACG 835
Db 809 CGCCCTAAACGATGCACTAGTGTGGGCAACTGATGCTTGATGCGTAGCTAACG 868
Qy 836 CGTGAAGTTGACCGGCTGGGAGTACGATCCGAATTTAAACCTAAGGAATTTAGCGG 895
Db 869 CGTGAAGTTGACCGGCTGGGAGTACGATCCGAATTTAAACCTAAGGAATTTAGCGG 928
Qy 896 GACCCGCACAAGCGGTGATTTATGATTAATTCGATGCAACGCGAAACCTTAACCTA 955
Db 929 GACCCGCACAAGCGGTGATGATGATGATTAATTCGATGCAACGCGAAACCTTAACCTG 988
Qy 956 CCCTTGACATGTAGCGAAATTTTCTAGAGATAGATTAGTG-CTTGGGAAACGCTAACAG 1014
Db 989 GTTTGACATGTAGCGAAATCTCCGAGACGAGAGAGATGCTTCCGGAGCCGTAAACAG 1048
Qy 1015 GTGCTGCATGCGTGTGCTCAAGCTCGTGTGAGATGTTGGGTTAAGTCCGCAACGAGC 1074
Db 1049 GTGCTGCATGCGTGTGCTCAAGCTCGTGTGAGATGTTGGGTTAAGTCCGCAACGAGC 1108
Qy 1075 GCACCCCTTGTCATTAATTCGATCATTTGTTGGGCACTTTAATAGACTCCCGGTAC 1134
Db 1109 GCACCCCTTGTCATTAATTCGATCATTTGTTGGGCACTTTAATAGACTCCCGGTAC 1168
Qy 1135 AAACCGAGAAAGGTGGGAGTACGTCATGCTTCATGAGCCCTTAATGGTAAAGGCTTAC 1194
Db 1169 AAACCGAGAAAGGTGGGAGTACGTCATGCTTCATGAGCCCTTAATGACAGGGCTTAC 1228
Qy 1195 ACGTAAATCAATGGGCGCTACAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAGAA 1254
Db 1229 ACGTAAATCAATGGTGTGACAGAGGTTGCCAACCCGAGGGGAGCTAATCTCAGAA 1288
Qy 1255 AGCGGCTGATGTCGAGATCGGATCTGCAACTCGACTCCGTAAGTCGSAATCGCTAGT 1314
Db 1289 AACCATGTGATCCGGATGCACTCTGCAACTCGAGTGCAGAGTGGAAATCGCTAGT 1348
Qy 1315 AATCGCGATGACATGTGCGCGGTAATCGTTCGCGGCTTGTACACACCGCCGTCA 1374
Db 1349 AATCGCGATGACATGTGCGCGGTAATCGTTCGCGGCTTGTACACACCGCCGTCA 1408
Qy 1375 CACCATGGAGTGGGTTTCCACAGAAAGAGTACTAATCCGTAGAGAGGGCGCTTGCA 1434
Db 1409 CACCATGGAGTGGGTTTCCACAGAAAGTAGGGTTAAACGCAAGAGTCCGCTTACCA 1468
Qy 1435 CGGTGAGATTCACTGAGGGGTG 1457
Db 1469 CGGTATGCTTCACTGAGCTGGGTG 1491

RESULT 44
US-10-029-397A-44
; Sequence 44, Application US/10029397A
; Publication No. US20030175709A1
; GENERAL INFORMATION:
; APPLICANT: MURPHY, GEORGE L.
; APPLICANT: WHITLEY, J. PENN
; TITLE OF INVENTION: METHOD AND SYSTEM FOR DEPLETING rRNA POPULATIONS
; FILE REFERENCE: AMI:076US
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 44
; LENGTH: 1544
```

```
TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-029-397A-44

Query Match      76.9%; Score 1119.8; DB 6; Length 1544;
Best Local Similarity 87.1%; Pred. No. 2.56-288;
Matches 1275; Conservative 0; Mismatches 182; Indels 6; Gaps 4;

Qy 1 ATTGAACGCTGGGGGCAATGCTTTACATGCAATGCAAGGGGAGCA--GGATGCTTGC 58
Db 29 ATTGAACGCTGGGGGCAATGCTTTACATGCAATGCAAGGGGAGCAAGGGAGCTTGC 88
Qy 59 ATCT--GGTGGCGATGGCGGACGGGTGATGATTCGAAACGTAATCCAGAAAGAGGGG 116
Db 89 TTCTGGGTGGCGATGGCGGACGGGTGATGATTCGAAACGTAATCCAGAAAGTACGGGTG 148
Qy 117 GGTAAACGATCGAAAGATGTGCTAATACCGCATTAATCTTAAGAGAGAAACAGGGATC 176
Db 149 GATTAATGATCGAAAGATGATGCTAATACCGCATTAATCTTAAGAGAGAAACAGGGAGC 208
Qy 177 GAAAGACCTTGGCGCTTTTGGAGCGGCGGATGCTGATTAAGTATGTTGGGTTAAAGC 236
Db 209 TTGGGCTTGGCGCTTATTCGAGCGGCGGATATGATTAAGTATGTTGGGTTAAAGC 268
Qy 237 CTACCAAGCGACGATGATGTTGCTCTGAGAGAGACGACGACACTGGGACTGAGA 296
Db 269 CCACCAAGCGACGATGATGTTGCTCTGAGAGAGATGATCCGCACTGGGACTGAGA 328
Qy 297 CACGGCCGACCTCTACCGGAGGACGACGATGGGAAATTTTGCATATGGGCGCAACCT 356
Db 329 CACGGCCGACCTCTACCGGAGGACGACGATGGGAAATTTTGCATATGGGCGCAACCT 388
Qy 357 GATCCAGCAATGCGCGGTGATGAGAAAGGCTTGGGTTGTAAGCTCTTTCAGTCGAG 416
Db 389 GATCCAGCAATGCGCGGTGATGAGAAAGGCTTGGGTTGTAAGCACTTTCAGTCGAG 448
Qy 417 AAGAAAGGTTACGTAATTAATGCTATGATGATGCAATGCAAGAGAAAGCACCGGC 476
Db 449 AAGAAAGGCTTGTGCAATTAATGCGCGCGGATGACGTAACCTGTAATAGCACCGGC 508
Qy 477 TAACTAGTGCACAGACCGCGGTAATACGTAAGGTGCAAGCTTAATCGAATTAATCGG 536
Db 509 TAACTAGTGCACACACCGCGGTAATACGTAAGGTGCAAGCTTAATCGAATTAATCGG 568
Qy 537 GCGTAAGGTTGGCGGACGCGCTTGTAAATGCAATGCAATGCCGCGCTTAACCTGG 596
Db 569 GCGTAAGGTTGGCGGACGCGCTTGTAAATGCAATGCAATGCCGCGCTTAACCTGG 628
Qy 597 GAATTCGCTTGAATTAACAAGCTGATGATGTCAGAGGAGGTGAAATTCATGTGTGA 656
Db 629 GAATTCGCTTGAATTAAGGTGATGTCAGAGTGTGTCAGAGGAGGTGAAATTCATGTGTGA 688
Qy 657 GCAGTGAATGCGTGAAGATATGGAAGAACATCGATGCGAAGGACGCTCTGGGTTAA 716
Db 689 GCAGTGAATGCGTGAAGATATGGAAGAACATCGATGCGAAGGACGCTCTGGGTTAA 748
Qy 717 CACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTTAATACCTGATGATGCA 776
Db 749 CACTGACGCTCATGACGAAAGCGTGGGAGCAACAGATTTAATACCTGATGATGCA 808
Qy 777 CGCCCTAAACGATGCACTAGTGTGGGCTTTATTA-GGCTTGTAACGAAGCTAACG 835
Db 809 CGCCCTAAACGATGCACTAGTGTGGGCAACTGATGCTTGATGCGTAGCTAACG 868
Qy 836 CGTGAAGTTGACCGGCTGGGAGTACGATCCGAATTTAAACCTAAGGAATTTAGCGG 895
Db 869 CGTGAAGTTGACCGGCTGGGAGTACGATCCGAATTTAAACCTAAGGAATTTAGCGG 928
Qy 896 GACCCGCACAAGCGGTGATTTATGATTAATTCGATGCAACGCGAAACCTTAACCTA 955
Db 929 GACCCGCACAAGCGGTGATGATGATTAATTCGATGCAACGCGAAACCTTAACCTG 988
Qy 956 CCCTTGACATGTAGCGAATTTTCTAGAGATAGATTAGTG-CTTGGGAAACGCTAACAG 1014
```

Db	989	GTTTTGACATGTGCGGAATCCTCCGGAACGAGAGATGCTTTGGAGCCGTAAACAG	1048
Qy	1015	GTGCGCAGATGGCTGTGCTGACAGTCTGTCGTGAGATGTTGGATTAAAGCCCGCAACAGC	1074
Db	1049	GTGCTGACATGGCTGTGCTGACAGTCTGTCGTGAGATGTTGGATTAAAGCCCGCAACAGC	1108
Qy	1075	GCAACCCCTTGTCAATTAATTTGCAATCTTTGGGCACTTAAATAGACTGCGCGTGAC	1134
Db	1109	GCAACCCCTTGTCAATTAATTTGCAATCTTTGGGCACTTAAATAGACTGCGCGTGAC	1168
Qy	1135	AAACCCGAGGAAGGTGGGGATGACGTCAAGTCCCATAGGCCCTTAATGGGTAGGGCTTAC	1194
Db	1169	AAACCCGAGGAAGGTGGGGATGACGTCAAGTCCCATAGGCCCTTAATGGGTAGGGCTTAC	1228
Qy	1195	ACGTAAATCAATAGCGCGCTAGACAGAGGTTTGCCAAACCCGAGAGGGGAGACTAATCTCAGA	1254
Db	1229	ACGTAAATCAATAGCGCGCTAGACAGAGGTTTGCCAAACCCGAGAGGGGAGACTAATCTCAGA	1288
Qy	1255	AGCGCGTGTAGTCCGGATCCGAGTCTTGCAACTCGACTCCGTGAAGTGGAAATGCTTAAGT	1314
Db	1289	AACCGATGTAGTCCGGATCCGAGTCTTGCAACTCGACTCCGTGAAGTGGAAATGCTTAAGT	1348
Qy	1315	AATGCCGATCAGCAATGTGCGGGTAATACGTTCCGGGCTCTTGACACACGCCCGCTCA	1374
Db	1349	AATGCCGATCAGCAATGTGCGGGTAATACGTTCCGGGCTCTTGACACACGCCCGCTCA	1408
Qy	1375	CACCAATGGGAGTGGGTTTCAACCAAGACAGGTAGTTAAACCGTAAGAAGGGCGCTTGCCA	1434
Db	1409	CACCAATGGGAGTGGGGAATACCAAGAAATAGGTAAACCGCAAGAGTCCGCTTAACA	1468
Qy	1435	CGGTGAATTCATGACTGGGGTGG	1457
Db	1469	CGGTATGCTTCAATGACTGGGGTGG	1491

```

RESULT 45
US-10-719-633-5
/ Sequence 5, Application US/10719633
/ Publication No. US20040137485A1
/ GENERAL INFORMATION:
/ APPLICANT: Iversen, Patrick L.
/ TITLE OF INVENTION: Antisense Antibacterial Method and
/ TITLE OF INVENTION: Composition
/ FILE REFERENCE: 0450-0032..30
/ CURRENT APPLICATION NUMBER: US/10/719,633
/ CURRENT FILING DATE: 2003-11-21
/ PRIOR APPLICATION NUMBER: US/09/726,774
/ PRIOR FILING DATE: 2000-11-29
/ PRIOR APPLICATION NUMBER: US 60/168,150
/ PRIOR FILING DATE: 1999-11-29
/ NUMBER OF SEQ ID NOS: 139
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 1544
/
/ TYPE: DNA
/
/ ORGANISM: Neisseria gonorrhoea
US-10-719-633-5

```

Query Match	76.9%;	Score 1119.8;	DB 7;	Length 1544;
Best Local Similarity	87.1%;	Pred. No. 2.5e-288;		
Matches 1275;	Conservative	0;	Mismatches 182;	Indels 6; Gaps 4
QY	1	ATTGAAACGTGGCGGCATGCTTTACACATGCAAGTCGAACGGCAGAC--GGATGCTTGC	58	
Db	29	ATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGACGGCAGCAGGAAAGCTTGC	88	
QY	59	ATCT--GGTGGCAGATGGCGGACGGGTGAGTAATGATCGGAACGTATCCAGAAAGGGG	116	
Db	89	TTCCTCGGGTGGCGAGTGGCGGAACGGGTGAGTAACATATCGGAACGTATCCGGTATGGGGGG	148	
QY	117	GGTAACGCATCGAAAGATGTGCTAATACCGCATATCTTAAAGAGGAAGCAGGGGATC	176	

Db	149	GATTAATCGATGAAAGAATCACTGTAATACCCGATACGCTTGTGAGAGGGAAGAAGCAAGGGGACC	208
Qy	177	GAAGAACCCTTCGCGCTTTTGGAGCGGCCGATGTCGTGATTAGCTAGTTGGTGGGTAAAGGC	236
Db	209	TTCCGGGCTTCGCTATCCGAGCGGCCGATATCTGATTAGCTGGTTGGCGGGTAAAGGC	268
Qy	237	CTACCAAGCGCAGCATGATGTTGTTGTCGAGAGGACGACGACGACAACTGGGACTTGAGA	296
Db	269	CCACCAAGCGCAGCATGATGTCGCGGGCTCTGAGAGGATGATCCGCAACTGGGACTAGA	328
Qy	297	CACGGCCCAAGCTCCTACGGGAGCAGAGTGGGAAATTTTGGACATGGGCGCAAGCCT	356
Db	329	CACGGCCCAAGCTCCTACGGGAGCAGAGTGGGAAATTTTGGACATGGGCGCAAGCCT	388
Qy	357	GATCAGCAATAGCGCGTGAAGTGAAGAAGCCTTCGGGTTGTAAAGCTTTTCAGTCGAG	416
Db	389	GATCAGCAATAGCGCGTGAAGTGAAGAAGCCTTCGGGTTGTAAAGCCTTTTCAGAGG	448
Qy	417	AAGAAAGGTTACGGTAAATTAATGTGACTCATGACGGTATCGACAGAGAAGACCGGC	476
Db	449	AAGAAAGGCTGTTTGCATATCGCGCGCCGATACGGTACTGAAAGATTAAGACCGGC	508
Qy	477	TAACTACGTCGACAGAGCGCGGTTAATCGTAAGGGTGCAGCGTTAATCGGAATTACG	536
Db	509	TAACTACGTCGACAGAGCGCGGTTAATCGTAAGGGTGCAGCGTTAATCGGAATTACG	568
Qy	537	GGGTAAAGGATGCGCAGCGCGCTTGTGAATGATGATGTAATCCCGGGCTTAAACCTGG	596
Db	569	GGGTAAAGGCGGCGCAGACGGTTACTTAAGCAGGATGTGAATCCCGGGCTCAACCGG	628
Qy	597	GAATTCGCTTTGAAACTACAAAGCTAGAGTGTGCAGAGGAGGTGCAATTCATGTCTA	656
Db	629	GAATTCGCTTTGAAACTAGAGTGTGCAGAGGAGGTGCAATTCATGTCTA	688
Qy	657	GCAATGAAATGCTGTGAATATGAGAAACATCAATGCGCAAGGACGCTCTGGGTTAA	716
Db	689	GCAATGAAATGCTGTGAATATGAGAAACATCAATGCGCAAGGACGCTCTGGGTTAA	748
Qy	717	CAGTACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAGTAACTCTGTAGTCCA	776
Db	749	CAGTACGCTCATGCTCCGAACCGTGGTACGAAACAGATTAGTAACTCTGTAGTCCA	808
Qy	777	CGCCCTAAACGATGTCAACTAGTTGTGGCCTTAATTAAGCTTGTGAACGATTAACG	835
Db	809	CGCCCTAAACGATGTCAACTAGTTGTGGCACTGATGCTTGTGTGCTGATGCTTAACG	868
Qy	836	CGTGAATGTAACCGCTGGGAGATACGCTCCGACAACTTAATACTCAAGAAATGACGGG	895
Db	869	CGTGAATGTAACCGCTGGGAGATACGCTCCGACAACTTAATACTCAAGAAATGACGGG	928
Qy	896	GACCGGCAACAGGGGTGATTATGTGGAATTAAATGATGCAACGCGAAGAACTTACCTA	955
Db	929	GACCGGCAACAGGGGTGATTATGTGGAATTAAATGATGCAACGCGAAGAACTTACCTG	988
Qy	956	CCCTTGACATGTAGCGAATTTTCTAGAGTAGATTAGTGCTTGGGAAACGTTAACACAG	1014
Db	989	GTTTGTGACATGTAGCGAATCTCTCCGAGACGAGAGTGTCTTGGGAGCGTTAACACAG	1048
Qy	1015	GTGCTGCAATGCTGTGCTGCTGCTGCTGTGCTGAGATGTTGGTTAAATGCTCCGCAAGAG	1074
Db	1049	GTGCTGCAATGCTGTGCTGCTGCTGCTGTGCTGAGATGTTGGTTAAATGCTCCGCAAGAG	1106
Qy	1075	GCAACCTTGTGATTAATGACATCAATTTGTTGGGCACTTAAATGAGACTGCGGTGAC	1133
Db	1109	GCAACCTTGTGATTAATGTTGTCATATTCGTTGGGCACTTAATGAGACTGCGGTGAC	1166
Qy	1135	AAACCGAGGAAGTGGGATGACGTCAGATCTCATGAGCCTTAATGGGTAGGCTTAC	1199
Db	1169	AAGCGAGGAAGTGGGATGACGTCAGATCTCATGAGCCTTAATGAGACTGCGGTGAC	1228
Qy	1195	ACGTAATTAATGAGCGGCTGACAGAGGTTGCCAACCCGAGAGGGGAGACTAATCTCAGAA	1255
Db	1229	ACGTAATTAATGAGCGGCTGACAGAGGTTGCCAACCCGAGAGGGGAGACTAATCTCAGAA	1288

```

Oy 1255 AGCGGTGTTAGTCCGATCGAGTCTGCACTCCGTAAGTGGAAATCGCTAGT 1314
    |||||
Db 1289 AACCGATGTTAGTCCGATGTTGCACTCTGCACTCGAAGTGGAAATCGCTAGT 1348
Oy 1315 AATCGCGATGAGTATGTCGCGGTAAATACGTTCCCGGGTCTTTGACACACCGCCGCA 1374
    |||||
Db 1349 AATCGCGATGAGTATGTCGCGGTAAATACGTTCCCGGGTCTTTGACACACCGCCGCA 1408
Oy 1375 CACCATGGAGTGGGTTTCCACCAAGACAGTATGTTAAACGTTAAAGAGGGCGTTGCCA 1434
    |||||
Db 1409 CACCATGGAGTGGGAGTATACCAAGATGATAGGATTAACCGCAAGAGTCCGCTTACCA 1468
Oy 1435 CGGTAGATTTCTGACTGGGGTG 1457
    |||||
Db 1469 CGGTATGCTTCATGACTGGGGTG 1491

RESULT 46
US-10-219-549-1
; Sequence 1, Application US/10219549
; Publication No. US20030170836A1
; GENERAL INFORMATION:
; APPLICANT: BRAMUCCI, MICHAEL G.
; APPLICANT: MCCUTCHEN, CAROL M.
; APPLICANT: MAGRAJAN, VASANTHA
; APPLICANT: THOMAS, STUART M.
; TITLE OF INVENTION: Microbial Production of Terephthalic Acid and
; FILE REFERENCE: CL-1257
; CURRENT APPLICATION NUMBER: US/10/219,549
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/091,645
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1451
; TYPE: DNA
; ORGANISM: Burkholderia sp.
US-10-219-549-1

Query Match 76.0%; Score 1108; DB 6; Length 1451;
Best Local Similarity 87.2%; Pred. No. 3.5e-285;
Matches 1552; Conservative 0; Mismatches 175; Indels 9; Gaps 3;

```

```

Oy 431 GTAAATATCTGACTCATGACGGTATCGACAGAAAGACACCGGCTTAACCTAGCCAG 490
    |||||
Db 431 GTTAATATCTGCGGGGATGACGGTATCGACAGAAATTAAGCACCGGCTTAACCTAGCCAG 489
Oy 491 CAGCCGCGGTAAATAGTATGAGTGCAGAGGTTAAATCGGAATTAAGGCGTAAAGGTCG 550
    |||||
Db 490 CAGCCGCGGTAAATAGTATGAGTGCAGAGGTTAAATCGGAATTAAGGCGTAAAGGTCG 549
Oy 551 CAGCGGCTTTGTAAGTCAATGTAATATCCCGGGCTTAACTGGGAATTCGTTGAA 610
    |||||
Db 550 CAGCGGCTTTGTAAGTCAATGTAATATCCCGGGCTTAACTGGGAATTCGTTGAA 609
Oy 611 ACTCAAGGCTTGAAGTATGAGGAGGAGGAGGGAATTCAGATGTAAGGAATGCGT 670
    |||||
Db 610 ACTCAAGGCTTGAAGTATGAGGAGGAGGAGGGAATTCAGATGTAAGGAATGCGT 669
Oy 671 AGAGATATGAAAGAACATCATGATGCGAAGGACGCTCTGGGTTAACTGACGCTCATG 730
    |||||
Db 670 AGAGATGTAAGGAATATCCGATGCGAAGGACGCTCTGGGCTTAATCTGACGCTCATG 729
Oy 731 CACGAAGCGTGGGAGCAACAGAGTTAGTATCCCTGTTAGTCAAGCCCTTAACGATG 790
    |||||
Db 730 CACGAAGCGTGGGAGCAACAGAGTTAGTATCCCTGTTAGTCAAGCCCTTAACGATG 789
Oy 791 TCAACTAGTTGTTGGGCTTTATAGGCTGGTAAAGCAAGCTTAAGCGGTAAGTTGACCGC 850
    |||||
Db 790 TCAACTAGTTGTTGGGCTTTATAGGCTGGTAAAGCAAGCTTAAGCGGTAAGTTGACCGC 849
Oy 851 CTGGGGAGTACGTCGCAAGATTAATACTCAAGAAATTAAGCGGGACCCGCAAGCGG 910
    |||||
Db 850 CTGGGGAGTACGTCGCAAGATTAATACTCAAGAAATTAAGCGGGACCCGCAAGCGG 909
Oy 911 TGAATATGATGATTAATTCGATGCAACCGGAAACCTTAACCTTACCTGACATGTAAC 970
    |||||
Db 910 TGAATATGATGATTAATTCGATGCAACCGGAAACCTTAACCTTACCTGACATGTAAC 969
Oy 971 GAATTTCTAGAGTATTAATTAATGAGTCT--TGGGAGCGCTTAACAGGTCGTGATGAGCT 1027
    |||||
Db 970 GAATTTCTAGAGTATTAATTAATGAGTCTCTGAAAGAAACGTAACAGGTCGTGATGAGCT 1029
Oy 1028 GTGCTACGCTGTCGTAAGATGTTGGGTTAACTCCCGCAACAGCGCAACCTTGTCA 1087
    |||||
Db 1030 GTGCTACGCTGTCGTAAGATGTTGGGTTAACTCCCGCAACAGCGCAACCTTGTCC 1089
Oy 1088 TTAATTCGATCAATTTGTTGGGCACTTAAATGAGTCTCCCGGTGACAAACCGAGAG 1147
    |||||
Db 1090 TTAATTCGATCAATTTGTTGGGCACTTAAATGAGTCTCCCGGTGACAAACCGAGAGAG 1144
Oy 1148 GTGGGATGAGTCAAGTCCCTCATAGGCTTAAATGAGGCTTCAACGTAATACATG 1207
    |||||
Db 1145 GTGGGATGAGTCAAGTCCCTCATAGGCTTAAATGAGGCTTCAACGTAATACATG 1204
Oy 1208 GCGGCTACAGAGGCTTCCCAACCGCGAGGAGGAGCTTAATCTCAAGAAAGCGCGTCTAGT 1267
    |||||
Db 1205 GTCGTAACAGAGGCTTCCCAACCGCGAGGAGGAGCTTAATCTCAAGAAAGCGCGTCTAGT 1264
Oy 1268 CCGGATGAGTCTGCAACTGACTCCGTGAAGTTCGGAATTCGTAATTCGCGGATCAG 1327
    |||||
Db 1265 CCGGATGAGTCTGCAACTGACTCCGTGAAGTTCGGAATTCGTAATTCGCGGATCAG 1324
Oy 1328 CATGTCGCGGTAATGAGTCCCGGGCTTGTACACACCGCGGCTCAACCATGGGAGTG 1387
    |||||
Db 1325 CATGTCGCGGTAATGAGTCCCGGGCTTGTGTACACACCGCGGCTCAACCATGGGAGTG 1384
Oy 1388 GGTTCACAGAAACAGATGATTAACGTTAAGAGGAGGCGCTTCCACGCGTGAAT 1443
    |||||
Db 1385 GGTTCGCAAGATGATGATTAACGTTAAGAGGAGGCGCTTTCACACGCGAGAT 1440

RESULT 47
US-10-464-724-4/c
; Sequence 4, Application US/10464724

```

```
; Publication No. US20030203398A1
; GENERAL INFORMATION:
; APPLICANT: Brumucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE REFERENCE: BC1033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,724
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 1453
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Denitrifying bacteria
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of rDNA corresponding to strain RA2
; US-10-464-724-4
Query Match      75.7%; Score 1103.2; DB 6; Length 1453;
Best Local Similarity 86.3%; Pred. No. 6,7e-284;
Matches 1258; Conservative 0; Mismatches 188; Indels 12; Gaps 3;

QY      1 ATTGAACGCTGGCGCGATCTTTACATGCAAGTGGAAACGACAGACGGATGCTTGAT 60
DB      1449 ATTGAACGCTGGCGCGATCTTTACATGCAAGTGGAAACGACAGCGCGCA---AC 1394
QY      61 CTGGTGGCGAGTGGCGCGAGCGGAGTGAATGATCCGAACGATCCAGAAAGGCGGCTA 120
DB      1393 CTGGCGCGAGCGCGCGAGCGGAGTGAATGATCCGAACGATCCAGAAAGTGGCGGAT 1334
QY      121 ACGCATGGAAGATGTGCTAATACCGCATATATCTTAAGAGAGAAAGAGGGATGAAA 180
DB      1333 GCCCGCGAAGAGCGCGATTAATATCCGATGTGATCTGAGGATGAAAGTGGCGGACCGCA 1274
QY      181 GACCTTCCGCTTTTGAAGCGCGCGATGTCTGATTACTAGTTGGTGGGTTAAAGGCTAC 240
DB      1273 GGGCTCACCGCTTTTGAAGCGCGCGATGTGATTAAGTGGTGGGTTAAAGGCTTAC 1214
QY      241 CAAGCGAGCATGATGATGTGTCTGAGAGAGAGCAACGACCACTGGGACTGAGACAG 300
DB      1213 CAAGCTGCGATTTGATGCTGTGTCTGAGAGAGATGATCAAGCACTGGGACTGAGACAG 1154
QY      301 GCCCAGACTCTTACGCGAGGAGCAGCAGTGGGGAATTTTGAACAATGGGCGCAAGCTGATC 360
DB      1153 GCCCAGACTCTTACGCGAGGAGCAGCAGTGGGGAATTTTGAACAATGGGCGCAAGCTGATC 1094
QY      361 CAGCAATGCCGCGTGTGAGTGAAGAGCGCTTCCGGTTGAAGCTCTTTCACTGAGAGA 420
DB      1093 CAGCCATTCGCGCTGTGAGAGAGAGCGCTTCCGGTTGAAGCTCTTTGTTAGGAGAG 1034
QY      421 AAAGTTACGTTAATATATGATCATGACGTTATCGACAGAGAGAGCAGCGGCTAAC 480
DB      1033 AAAGGCTCTCTTAAATACAGGGGCGATATACGTTACCGTAAGATTAAGACCGGCTAAC 974
QY      481 TACGTCCAGCAGACCGCGGTAATATAGTGGTCAAGCGTTAATCGGAATTACTGGGCGT 540
DB      973 TACGTCCAGCAGACCGCGGTAATATAGTGGTCAAGCGTTAATCGGAATTACTGGGCGT 914
QY      541 AAAGGTGGCGAGCGCGCTTTTGAAGTCAATGTGAATATCCCGGGCTTAACTGGGAAT 600
DB      913 AAAGCGTGGCGAGCGCGCTTTTGAAGTCAAGTGAATATCCCGGGCTTCAACTGGGAAC 854
QY      601 TGGCTTTGAAGTCAAGAGCTAGAGTGTGGCAGAGGAGGTGAATTCATGTGTAGCAG 660
DB      853 TGGCTTTGTAGCTGCAAGGCTGAGAGTGGCGCAGAGGAGGTGAATTCCTCGGTTAGCAG 794
QY      661 TGAATTCGTAAGATATGAGAAAGATCGATGCGAAGGACGCTTCCGGTTAACA 720
```

```
DB      793 TGAATTCGTAAGATATGAGAAAGATCGATGCGAAGGACATCCCTGGGCGTGCAC 734
QY      721 GACGCTATGACAGAAACGTTGGGAGCAAAACAGATTAGTACCTGTTAGTCCACGCC 780
DB      723 GACGCTATGACAGAAACGTTGGGAGCAAAACAGATTAGTACCTGTTAGTCCACGCC 674
QY      781 CTAAACGATGTCACTAGTTTGGGCGCTTTTGAAGGCTGTGAACGAGCTAACGCTGA 840
DB      673 CTAAACGATGTCACTAGTTTGGGAGTTCAATTTTCAAGTACGAAAGCTTAACGCTGA 614
QY      841 AGTTACCGCTTGGGAGTACGTTGCAAGATTAACTCAAGAGATTGAACGGGAGCC 900
DB      613 AGTTACCGCTTGGGAGTACGTTGCAAGATTAACTCAAGAGATTGAACGGGAGCC 554
QY      901 GCACAGCGGTGATATGATGATTAATTCATGCAACGCGAAAACCTTACCTACCTT 960
DB      553 GCACAGCGGTGATGATGATGATTAATTCATGCAACGCGAAAACCTTACCTACCTT 494
QY      961 GACATGTACGAAATTTTCTAGAGATGATAGTGTCT---TGGGAAACGTTAACAGAGTG 1017
DB      493 GACATGTACGAAATCTGCAAGAGACAGAGAGTGTCTGAAAGAGACGTTAACAGAGTG 434
QY      1018 CTGCAATGGCTGTGCTCAGCTGCTGCTGAGATGTTGGTTAAGTCCGCAACGACGCA 1077
DB      433 CTGCAATGGCTGTGCTCAGCTGCTGCTGAGATGTTGGTTAAGTCCGCAACGACGCA 374
QY      1078 ACCCTTGCATTAATATTCGATCATATTTGTTGGGACCTTAAATGAGATGTCGGGTGACAA 1137
DB      373 ACCCTTGCATTAATATTCGATCATATTTGTTGGGACCTTAAATGAGATGTCGGGTGACAA 319
QY      1138 CCGAGAGAAAGTGGGAGTGAAGTCAAGTCTTATGAGGCTTTATGGTGGGCTTACAG 1197
DB      318 CCGAGAGAAAGTGGGAGTGAAGTCAAGTCTTATGAGGCTTTATGGTGGGCTTACAG 259
QY      1198 TAAATCAATGGCGCGTACAGAGGTTGGCCAAACCCGAGAGGGAGCTAATCTCAGAAAC 1257
DB      258 TAAATCAATGGCGCGTACAGAGGTTGGCCAAACCCGAGAGGGAGCTAATCTCAGAAAC 199
QY      1258 GCGTGTAGTCCGATCCGATCCGATCTGCAACTGACTCCGTAAGTGGAAATGCTAGTAAT 1317
DB      198 GCGTGTAGTCCGATCCGATCCGATCTGCAACTGACTCCGTAAGTGGAAATGCTAGTAAT 139
QY      1318 CCGGATTCAGATGTGCGGTTGAATACGTTCCCGGCTTTTGAACACACCGCCGTCAC 1377
DB      138 CCGGATTCAGATGTGCGGTTGAATACGTTCCCGGCTTTTGAACACACCGCCGTCAC 79
QY      1378 CATGGAGTGGGTTTCAACAGAGAGGATGCTAAACGTTAAGAGAGGCGCTTGCACAG 1437
DB      78 CATGGAGTGGGTTTCCAGAGAGGATGTTAGCTTAACCCAGAGAGGCGGATTAACACAG 19
QY      1438 TGAGATTCACTGACTGGG 1455
DB      18 CAGGCTTCGTACTGGG 1

RESULT 48
US-10-464-356-4/c
; Sequence 4, Application US/10464356
; Publication No. US20030207320A1
; GENERAL INFORMATION:
; APPLICANT: Brumucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE REFERENCE: BC1033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,356
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
```

SEQ ID NO 4  
LENGTH: 1453  
TYPE: DNA  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Dinitrifying bacteria  
FEATURE:  
OTHER INFORMATION: Nucleotide sequence of rDNA corresponding to strain RA2  
US-10-464-356-4

Query Match 75.7%; Score 1103.2; DB 6; Length 1453;  
Best Local Similarity 86.3%; Pred. No. 6.7e-284;  
Matches 1258; Conservative 0; Mismatches 188; Indels 12; Gaps 3;

1 ATTGAACGCTGGCGGCACTTTTACACATGCAAGTCGAAACGGCAGCAGGATGCTTCAT 60  
1449 ATTGAACGCTGGCGGCACTTTTACACATGCAAGTCGAAACGGCAGCAGGATGCTTCAT 1394  
61 CTGGTGGGAGTGGGCGGACGGGTGAGTATGATCGGAACTATCGAAGAGGGGGTGA 120  
1393 CTGGGCGGAGCGCGGAAAGGGGTGAGTATGATCGGAACTATCGAAGAGGGGGTGA 1334  
121 ACGCATCGAAGAGTGTCTAATACCGCATATATCTTAAGAGAGAAAGCGGATCGAA 180  
1333 GCCCGCGAAAGCCGGAATTAACCGCATGTATCTGAGATGAAAGTGGGGGACCGCAA 1274  
181 GACCTTGGGCTTTTGGAGCGGCGCATGTCTGATTAAGTGTGGTGGTAAAGGCTTAC 240  
1273 GCCCTCAGCGCTTTGAGCGCGCATGTGCGAATTAAGTGTGGGTTAAAGGCTTAC 1214  
241 CAAGCGCAGCATGATGTTGGTCTGAGAGAGACCAAGCCACATGGGACTGAGACAG 300  
1213 CAAGCTGCGATCTTGAAGTGTCTGAGAGAGATGATCAAGCCACATGGGACTGAGACAG 1154  
301 GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACATAGGGCGCAAGCTGATC 360  
1153 GCCCAGACTCTTACCGGAGGAGCAGAGTGGGGAATTTTGGACATAGGGCGCAAGCTGATC 1094  
361 CAGCATGCGCGGCTGAGAGAGAGGCTTGGGTTTGAAGCTCTTTCACTGAGAGAGA 420  
1093 CAGCATGCGCGGCTGAGAGAGAGGCTTGGGTTTGAAGCTCTTTCACTGAGAGAGA 1034  
421 AAAGGTTACGTTAATTCGTACTGACGATGACGATGACGAGAGAGAGACCGGCTAAC 480  
1033 AAAGGTTACGTTAATTCGTACTGACGATGACGATGACGAGAGAGAGACCGGCTAAC 974  
481 TACGTGCGAGCAGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGGAATTAAGTGGCGT 540  
973 TACGTGCGAGCAGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGGAATTAAGTGGCGT 914  
541 AAAGGTTGCGAGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGGAATTAAGTGGCGT 600  
913 AAAGGTTGCGAGCGCGGTAAATACGTAGGGTGAAGCGTTAATCGGAATTAAGTGGCGT 854  
601 TCGGTTGAAACTACAGAGCTGAGTGGGAGGAGGAGTGAATTCATGTAGAGAG 660  
853 TCGGTTGAAACTACAGAGCTGAGTGGGAGGAGGAGTGAATTCATGTAGAGAG 794  
661 TGAATGCGTGAAGATATGAGAGACATCGATGGCGAAGGCAAGCTCTCGGTTAACT 720  
793 TGAATGCGTGAAGATATGAGAGACATCGATGGCGAAGGCAAGCTCTCGGTTAACT 734  
721 GACGCTCATGCAAGAAAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
733 GACGCTCATGCAAGAAAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 674  
781 CTAAGCATGTCATTAAGTGTGGGCTTATTAAGCTTGGTGGTGAAGGAGGAGGAGGAGGAG 840  
673 CTAAGCATGTCATTAAGTGTGGGCTTATTAAGCTTGGTGGTGAAGGAGGAGGAGGAGGAG 614  
841 AGTTGACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
613 AGTTGACCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 554

901 GCAACAGCGGTGATATGATGATTAATTCGATGCAACGCGAAGAAACCTTACTACCTT 960  
553 GCAACAGCGGTGATATGATGATTAATTCGATGCAACGCGAAGAAACCTTACTACCTT 494  
961 GACATGTAGCGAATTTTCTAGATAGATTAAGTGTCT--TCGGAAAGCTTAACAGAGTG 1017  
493 GACATGTAGCGAATTTTCTAGATAGATTAAGTGTCT--TCGGAAAGCTTAACAGAGTG 434  
1018 CTGCATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
433 CTGCATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374  
1078 ACCCTTGCATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
373 ACCCTTGCATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319  
1138 CCGAGGAAAGGTGGGAGTGAAGTCAAGTCTCATGCTTATGAGTGGGCTTACACAG 1197  
318 CCGAGGAAAGGTGGGAGTGAAGTCAAGTCTCATGCTTATGAGTGGGCTTACACAG 259  
1198 TAAATCAATGCGCGGTACAGAGGTTGCTCAACCGCGAGGGGAGCTTAATTCAGAAAGC 1257  
258 TCATACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 199  
1258 GCGTGTAGTCCGAGTCCGAGTCTGCAACTGCACTGCTGCAAGTCCGAAATGCTTAAAT 1317  
198 CAGTGTAGTCCGAGTCCGAGTCTGCAACTGCACTGCTGCAAGTCCGAAATGCTTAAAT 139  
1318 CCGGATCAGCATGTCCGAGTGAATTCGCTTCCGAGTCTTGTACACACCGCCCTGACAC 1377  
138 CCGGATCAGCATGTCCGAGTGAATTCGCTTCCGAGTCTTGTACACACCGCCCTGACAC 79  
1378 CATGAGAGTGGGTTTACCAAGAGGATGATCTTAACGTTAAGAGGAGGCTTTCGACAGG 1437  
78 CATGAGAGTGGGTTTACCAAGAGGATGATCTTAACGTTAAGAGGAGGCTTTCGACAGG 19  
1438 TGAGATTCAGTCTGGG 1455  
18 CAGGTTCTGATCTGGG 1

RESULT 49  
US-10-464-709-4/C  
Sequence 4, Application US/10464709  
Publication No. US20030207321A1  
GENERAL INFORMATION:  
APPLICANT: Brumucci, Michael  
TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
TITLE OF INVENTION: Industrial Wastewater Bioreactors  
FILE REFERENCE: BCI03 US NA  
CURRENT FILING DATE: 2003-06-17  
PRIOR APPLICATION NUMBER: US/09/735,567  
PRIOR FILING DATE: 2000-12-13  
PRIOR APPLICATION NUMBER: 60/117,140  
PRIOR FILING DATE: 16 DECEMBER 1999  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 4  
LENGTH: 1453  
TYPE: DNA  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Dinitrifying bacteria  
FEATURE:  
OTHER INFORMATION: Nucleotide sequence of rDNA corresponding to strain RA2  
US-10-464-709-4

Query Match 75.7%; Score 1103.2; DB 6; Length 1453;  
Best Local Similarity 86.3%; Pred. No. 6.7e-284;  
Matches 1258; Conservative 0; Mismatches 188; Indels 12; Gaps 3;

1 ATTGAACGTGGCGGCATCTTTAACAATGCAAGTCGAAACGGCAGACACGGATCTTGCAAT 60  
1449 ATTGAACGTGGCGGCATCTTTAACAATGCAAGTCGAAACGGCAGCGCGGCA-----AC 1394  
61 CTGGTGGCGAGTGGCGGACGGGTGAGTAAATGCAATCGGAAACGTAACAGAAAGAGGGGGGTA 120  
1393 CTGGCGGCGAGCGGCGGAAAGGGTGAATACATCGGAAACGTCGACGAGCGTGGGGGATTA 1334  
121 ACGCATCGAAAGATGCTAATACCGCATATCTCTAAGAGAGAAAGAGGGGATCGAAA 180  
1333 GCCCGGCGAAACCGGATTAATACCGCATATCTGATCTGAGGATGAAAGTGGGGGACCGCAA 1274  
181 GACCTTGGCTTTTGGAGCGCGCATGTCTGATTAGCTGATTGTTGGGTAAAGGCTTAC 240  
1273 GGCTTCAGCGCTTTGGAGCGCGCATGTGAGATTAGTGTGGGTAAAGGCTTAC 1214  
241 CAAGCGACGATCAATGATTGTTGTTGAGAGAGACACACACCACTGGGAGCTGAGACAG 300  
1213 CAAGCTTGCGATCTGATGCTGTTGAGAGGATGATCAGCCACACTGGGACTGAGACAG 1154  
301 GCCCGACCTCTTACCGGAGGACAGCAGTGGGGAATTTTGGACAATGCGCGCAAGCTTATC 360  
1153 GCCCAGACTCTTACCGGAGGACAGCAGTGGGGAATTTTGGACAATGCGCGCAAGCTTATC 1094  
361 CAGCAATGCGCGCTGATGAGAAAGGCTTGGGTTGTAAGCTCTTTCAGTTCGAGAAAG 420  
1093 CAGCATTTCGCGGTGACGAGACGAAAGGCTTGGGTTGTAAGCTCTTTCAGTTCGAGAAAG 1034  
421 AAAGTTACGGTAATTAATCTGTAATCTGATGACGCTTCCAGACAGAAAGACCGGCTAAC 480  
1033 AAGGGCTCTCTCTTAATACAGGGGCGCATATGACGCTACCGTAAATTAAGACCGGCTAAC 974  
481 TACGTGCGACAGCGCGCGGTAATACGTAAGGTCGAAAGCGCTTATCGGAATTAATCTGGCGCT 540  
973 TACGTGCGACAGCGCGCGGTAATACGTAAGGTCGAAAGCGCTTATCGGAATTAATCTGGCGCT 914  
541 AAAGGTGCGACAGCGCGCTTGTAGTCAATGTAATCCCGGGCTTAACTCTGGGAAT 600  
913 AAAGGTGCGACAGCGCGCTTGTAGTCAATGTAATCCCGGGCTTAACTCTGGGAAT 854  
601 TGGCTTTGAAATCTAACAAGCTTAAGTGTGCGACAGAGGATGGAATTCATGTTGACAG 660  
853 TGGCTTTGAAATCTAACAAGCTTAAGTGTGCGACAGAGGATGGAATTCATGTTGACAG 794  
661 TGAATGCGTGAATATGGAAGAATCGATGCGGACGAGCTCTGGGTTAACTCT 720  
793 TGAATGCGTGAATATGGAAGAATCGATGCGGACGAGCTCTGGGTTAACTCT 734  
721 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGTACCTGATGATCCAGCC 780  
733 GACGCTCATGACGAAAGCGTGGGAGCAACAGATTAAGTACCTGATGATCCAGCC 674  
781 CTAAACGATGTAATCTAAGTGTGGGCTTAAATGCTTGTGTAACGAAGCTTAACGCTGA 840  
673 CTAAACGATGTAATCTAAGTGTGGGCTTAAATGCTTGTGTAACGAAGCTTAACGCTGA 614  
841 AGTTGACCGCTGGGAGGTAACGTCGCAAGATTAACTCAAGAAATTAAGAGGAGGAGCC 900  
613 AGTTGACCGCTGGGAGGTAACGTCGCAAGATTAACTCAAGAAATTAAGAGGAGGAGCC 554  
901 GCAACAGCGGTGATTAATGATTAATGATGCAACGCGGAAACCTTAACCTACCTT 960  
553 GCAACAGCGGTGATTAATGATTAATGATGCAACGCGGAAACCTTAACCTACCTT 494  
961 GACATGTAAGGAATTTTCTAGATTAATGATGCT--TCGGGAGCGTAAACAGAGT 1017  
493 GACATGTAAGGAATTTTCTAGATTAATGATGCT--TCGGGAGCGTAAACAGAGT 434  
1018 CTGATAGGCTGTGCTGAGCTGCTGATGAGATGTTGGGTTAACTCCGCGACAGAGGCA 1077  
433 CTGATAGGCTGTGCTGAGCTGCTGATGAGATGTTGGGTTAACTCCGCGACAGAGGCA 374  
1078 ACCCTGTCAATTAATGCAATCATTTGTTGGGCACTTAAATGAGACTGCGGATGACAAA 1137

DB 373 ACCCTGTCAATTAATGCTTACGAAA-----GGGCACTCTAATGGAAGTGGCGGTACAAA 319  
1138 CCGAGGAAGGTGGGAGTGAAGTCAAGTCTTCATGCGCCCTTATGGGTAGGGCTTACACAG 1197  
318 CCGAGGAAGGTGGGAGTGAAGTCAAGTCTTCATGCGCCCTTATGGGTAGGGCTTACACAG 259  
1138 TAAATCAATGCGGCTTACAGAGGCTTTCGCAACCGCGAGGGGGAGCTAATCTCAGAAAGC 1257  
258 TCATCAATGCGGCTTACAGAGGCTTTCGCAACCGCGAGGGGGAGCTAATCTCAGAAAGC 199  
1258 GCGTGTAGTCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 1317  
138 CAGTGTGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 139  
1318 CCGGATGACAGATGTCGCGGTAATACGTTCCCGGCTTGTACACACCGCCCTTACACAG 1377  
138 CCGGATGACAGATGTCGCGGTAATACGTTCCCGGCTTGTACACACCGCCCTTACACAG 79  
1378 CATGGAGTGGGTTTCAACAGAGAGGATGTAACGTTAAGGAGGGGCGCTTGCACAG 1437  
78 CATGGAGTGGGTTTCAACAGAGAGGATGTAACGTTAAGGAGGGGCGATTAACACAG 19  
1438 TGAGATTCAATGACGAGG 1455  
DB 18 CAGGCTTCTGATCTGGG 1

## RESULT 50

US-10-678-023-1  
; Sequence 1, Application US/10678023  
; Publication No. US20050074431A1  
; GENERAL INFORMATION:  
; APPLICANT: Martin et al., Phyllis A. W.  
; TITLE OF INVENTION: CHROMOBACTERIUM SUTTSUGA SP. NOV. AND USE FOR CONTROL OF INSECT  
; TITLE OF INVENTION: PESTS  
; FILE REFERENCE: 0136.02  
; CURRENT APPLICATION NUMBER: US/10/678,023  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 1440  
; TYPE: DNA  
; ORGANISM: Chromobacterium suttsuga NRRL B-30655  
US-10-678-023-1

Query Match 75.7%; Score 1103; DB 9; Length 1440;  
Best Local Similarity 87.5%; Pred. No. 7.5e-284;  
Matches 1241; Conservative 1; Mismatches 172; Indels 5; Gaps 3;

1 ATTGAACGTGGCGGCATCTTTAACAATGCAAGTCGAAACGGCAGACCGGATCTTGCAAT 60  
24 ATTGAACGTGGCGGCATCTTTAACAATGCAAGTCGAAACGGCAGACCGGATCTTGCAAT 82  
61 CTGGTGGCGAGTGGCGGACGGGTGAGTAAATGCAATCGGAAACGTAACAGAAAGAGGGGGTA 120  
83 C-GCTGACGAGTGGCGGACGGGTGAGTAAATGCTCGAAATGACCTGTATATGAGGAGTA 141  
121 ACGCATCGAAAGATGCTAATACCGCATATCTTAAGAGAGAAAGAGGGGATCGAAA 180  
142 GCTGCGGAAAGCGGATTAATACCGCATATCTTAAGAGAGAAAGAGGGGATCGAAA 201  
181 GACCTTGGCTTTTGGAGCGCGCATGTCTGATTAGCTTATGTTGGGTAAAGGCTTAC 240  
202 GACCTTGGCTTTTGGAGCGCGCATGTCTGATTAGCTTATGTTGGGTAAAGGCTTAC 261  
241 CAAGCGACGATCAATGATTGTTGTTGAGAGAGACAGACCACTGGGACTGAGACAG 300  
262 CAAGCGACGATCAATGATTGTTGTTGAGAGAGATCCGACACTGGGACTGAGACAG 321  
301 GCCCAGACTCTTACCGGAGGACAGTGGGGAATTTTGGACAATGAGGCGCAAGCTTATC 360

Db 322 GCCCAGACTCCTACGGAGGCGACGATGGGGAAATTTTGGACATGGGGCAACCTGATC 381  
Qy 361 CAGCAATGCCCGGTGATGAGAGAGCCTTCCGGGTGTTAAAGCTCTTTCAGTGAGAA 420  
Db 382 CAGCATCCCGGTGTCTGAGAGAGAGCCTTCCGGGTGTTAAAGACTTTTGTCCAGGAG 441  
Qy 421 AAAGGTTACGGTAATTAATCGTACTCATGACGATGATGACAGAGAGAGAGAGAGAG 480  
Db 442 AATCCCGGTGTTAAATACCCGGGGGATGACATGACBAGAGATTAAGACACCGCTAAC 501  
Qy 481 TACGTCCAGCAGCCCGGTATATACGTAGGGTGAAGCGTTATCGAATTAATCTGGCGT 540  
Db 502 TACGTCCAGCAGCCCGGTATATACGTAGGGTGAAGCGTTATCGAATTAATCTGGCGT 561  
Qy 541 AAAGGTGGCAGCGCGCTTTGTAAGTCAAGATGTGAATCCCGGGCTTAACCTGGGAAT 600  
Db 562 AAAGGTGGCAGCGCGCTTTGTAAGTCAAGATGTGAATCCCGGGCTTAACCTGGGAAT 621  
Qy 601 TCCGTTTGAATCTAAGGCTAGAGTGTGGCAGAGGGAGTGGAAATTCATGTGTAGAG 660  
Db 622 GGCATTGAGACTGCAAGCTAGAGTGTGGCAGAGGGAGTGGAAATTCATGTGTAGAG 681  
Qy 661 TGAATGCGTAGAGATATGAGAAACATCGATGGCGAAGGACCTCTGCGTTAACT 720  
Db 682 TGAATGCGTAGAGATATGAGAAACATCGATGGCGAAGGACCTCTGCGTTAACT 741  
Qy 721 GACGCTATGACGAGAAAGCGTGGGAGCAACAGATTAATCCTGTGTAGTCCAGCGC 780  
Db 742 GACGCTATGACGAGAAAGCGTGGGAGCAACAGATTAATCCTGTGTAGTCCAGCGC 801  
Qy 781 CTAAACGATGCAATAGTGTGGGCGCTTATTAAGCTTGTGTAAGAGCTAACCGGTGA 840  
Db 802 CTAAACGATGCAATAGTGTGGGCGCTTATTAAGCTTGTGTAAGAGCTAACCGGTGA 861  
Qy 841 AGTTGACCCGCTGGGAGATACGCTGCAAGATTAATACTCAAGAGATTAAGAGGAGCC 900  
Db 862 AATTGACCCGCTGGGAGATACGCTGCAAGATTAATACTCAAGAGATTAAGAGGAGCC 921  
Qy 901 GCACAGAGCGGTGATATATGATTAATTCGATGCAAGCGGAAACCTTACCTTACCTT 960  
Db 922 GCACAGAGCGGTGATATATGATTAATTCGATGCAAGCGGAAACCTTACCTTACCTT 981  
Qy 961 GACATGTAGCGAATTTTCTAGAGATAGATAGTGC---TTCGGGAACCTTAACAGAGTG 1017  
Db 982 GACATGTAGCGAATTTTCTGAGATATCTTGTGTGCCCAAGAGAGCGGTAAACAGAGTG 1041  
Qy 1018 CTGCAATGCTGTCTGACCTGTGTCTGAGATGTTGGGTTAAGTCCGCAACGAGCGCA 1077  
Db 1042 CTGCAATGCTGTCTGACCTGTGTCTGAGATGTTGGGTTAAGTCCGCAACGAGCGCA 1101  
Qy 1078 ACCCTTGTCAATTAATGCTATATTTGGTGGGCACTTAAATGACATGCGGTGACAA 1137  
Db 1102 ACCCTTGTCAATTAATGCTATATTTGGGCACTTAAATGACATGCGGTGACAA 1161  
Qy 1138 CCGAGAGAGGTGGGATGACGTCAGATCTCATGACCTTAATGGGTAGAGCTTACAG 1197  
Db 1162 CCGAGAGAGGTGGGATGACGTCAGATCTCATGATGACCTTAATGAGAGAGCTTACAG 1221  
Qy 1198 TAATACAAATGCGCTACAGAGGTTGCCAACCCCGAGAGGGGAGCTAATCTCAGAAAG 1257  
Db 1222 TCATACAAATGCGCTACAGAGGTTGCCAACCCCGAGAGGGTGTCCAAATCTCATTAAC 1281  
Qy 1258 GCGTGTATGTCGGATGCGAGATCTGCAACTGACTCCGTGAAGTGGAAATGCTAGTAAT 1317  
Db 1282 CGATGTGTATGTCGGATGCGAGATCTGCAACTGAGTGCATGAGTGGAAATGCTAGTAAT 1341  
Qy 1318 CCGGATCAGCATGTCGGGATGAGTAATGCTCCGGGTCTTGTACACACGCGCGTACAGC 1377  
Db 1342 CCGGATCAGCATGTCGGGATGAGTAATGCTCCGGGTCTTGTACACACGCGCGTACAGC 1401  
Qy 1378 CATGGAGTGGATTTCACACAGAGAGGTAGTCTAACCG 1416  
Db 1402 CATGGAGTGGATTTCACACAGAGAGGTAGTCTAACCG 1440

RESULT 51  
US-10-915-740A-8  
; Sequence 8, Application US/10915740A  
; Publication No. US20050191316M1  
; GENERAL INFORMATION:  
; APPLICANT: Frazer, Claire M.  
; APPLICANT: Hickey, Erin  
; APPLICANT: Peterson, Jeremy  
; APPLICANT: Tectellin, Hervé  
; APPLICANT: Ventier, J. Craig  
; APPLICANT: Masignani, Vega  
; APPLICANT: Galeotti, Cealira  
; APPLICANT: Mora, Mantova  
; APPLICANT: Ratti, Giulio  
; APPLICANT: Scarselli, Maria  
; APPLICANT: Scarlato, Vincenzo  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizze, Mariagrazia  
; APPLICANT: Grandi, Guido  
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use  
; FILE REFERENCE: 002441.00090  
; CURRENT FILING DATE: US/10/915,740A  
; PRIOR FILING DATE: 2004-08-11  
; PRIOR APPLICATION NUMBER: 09/806,866  
; PRIOR FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: USN 60/103,794  
; PRIOR FILING DATE: 1998-10-09  
; PRIOR APPLICATION NUMBER: USN 60/132,068  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: PCT/US99/25373  
; PRIOR FILING DATE: 1999-10-08  
; NUMBER OF SEQ ID NOS: 1068  
; SOFTWARE: Patent In version 3.2  
; SEQ ID NO 8  
; LENGTH: 20844  
; TYPE: DNA  
; ORGANISM: Neisseria meningitidis  
US-10-915-740A-8  
Query Match 75.5%; Score 1100.6; DB 9; Length 20844;  
Best Local Similarity 87.2%; Pred. No. 5.1e-283;  
Matches 1276; Conservative 0; Mismatches 179; Indels 8; Gaps 6;  
Qy 1 ATGGAAGCTGCGCGCATGCTTTACACATGCAATGCCAGCGCAC--GGATGCTTGC 58  
Db 18939 ATTGAACGCTGCGCGCATGCTTTACACATGCAATGCCAGCGCACAGAGAGCTTGC 18998  
Qy 59 ATCT--GGTGGCGATGGCGGAGCGGGTGAATGCAATCGGAACGTATCCAGAGAGGG 116  
Db 18999 TTCTGGGTGGCGATGGCGGAGCGGGTGAATGCAATCGGAACGTATCCAGAGAGGG 19058  
Qy 117 GGTAAAGCATGAAAGATGTGCTAATACCGCATTAATCTAAGAGAGAAAGCAGGGATC 176  
Db 19059 GATTAACGATGAAAGATGATGCTAATACCGCATTAATCTAAGAGAGAAAGCAGGGATC 19118  
Qy 177 GAAAGACCTTGGCGCTTTTGAAGCGCGCGATGTCTGATTAAGTGTGGGTAAAGGC 236  
Db 19119 TTCGGGCTTGGCGCTTTTGAAGCGCGCGATGTCTGATTAAGTGTGGGTAAAGGC 19178  
Qy 237 CTACCAAGGCGAGATAGTATGTTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 296  
Db 19179 CTACCAAGGCGAGATAGTATGTTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19238  
Qy 297 CACGCGCAGACTCTTACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 356  
Db 19239 CACGCGCAGACTCTTACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19297  
Qy 357 GATCAGAGATGCGCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416  
Db 19298 GATCAGAGATGCGCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19357



```
QY 417 AAGAAAGTTACGGTAATATCTGACTGACGCTGATGACAGAAAGACCGGC 476
DB 19358 AAGAAAGCTGTGTATATATACAGCGGCTGATGACGTAACCTGAAGATTAAGACCGGC 19417
QY 477 TAACTACGTCAGACGACCGCGGTAACTAGTGGGTGAAGGTTAATCGGAATTAACGG 536
DB 19418 TAACTACGTCAGACGACCGCGGTAACTAGTGGGTGAAGGTTAATCGGAATTAACGG 19476
QY 537 GCGTAAAGGAGTGCAGACGCGGCTTGTAAAGTCAAGTGTGAAATCCCGGCGCTTAACTGG 596
DB 19477 GCGTAAAGGAGTGCAGACGCGGCTTGTAAAGTCAAGTGTGAAATCCCGGCGCTTAACTGG 19536
QY 597 GAATTCGCTTTGAAACTACAAAGCTAAGATGTGGCAGAGGAGGTGGAATTCATGTGTGA 656
DB 19537 GAATTCGCTTTGAAACTACAAAGCTAAGATGTGGCAGAGGAGGTGGAATTCATGTGTGA 19596
QY 657 GCAATGAAATGCGTAGAGATATGGAAGAACTGATGCGGAAGGCGCTCCGCGGTTAA 716
DB 19597 GCAATGAAATGCGTAGAGATATGGAAGAACTGATGCGGAAGGCGCTCCGCGGTTAA 19656
QY 717 CACTGACGCTACGACGAAAGCGTGGGAGCAAAAGAGATTAGATACCTGGTAGTCCA 776
DB 19657 CACTGACGCTACGACGAAAGCGTGGGAGCAAAAGAGATTAGATACCTGGTAGTCCA 19716
QY 777 CCGCCCTAAACGATGTCACACTAGTTGTGGGCTTATTA-GGCTTGTAAAGAACTAACG 835
DB 19717 CCGCCCTAAACGATGTCACACTAGTTGTGGGCTTATTA-GGCTTGTAAAGAACTAACG 19776
QY 836 CGTGAAGTTGACCGGCTGGGAGATAGGTGCGAAAGTTAAACTCAAAAGAAATGACGGG 895
DB 19777 CGTGAAGTTGACCGGCTGGGAGATAGGTGCGAAAGTTAAACTCAAAAGAAATGACGGG 19836
QY 896 GAACCGCACAAGGCGGTGATATGATGATTAATTCGATGCAACGCGAAAACTTACTCTA 955
DB 19837 GAACCGCACAAGGCGGTGATATGATGATTAATTCGATGCAACGCGAAAACTTACTCTA 19896
QY 956 CCCTTGACATGTAGCGAATTTTGTAGAGATGATTAAGTGTGCGGAACGCTTAACAG 1014
DB 19897 CCCTTGACATGTAGCGAATTTTGTAGAGATGATTAAGTGTGCGGAACGCTTAACAG 19956
QY 1015 GTGCTGATGAGCTGTGTGCTGCTGTGTGCTGATGATGTTGGTTAGTCCCGGACGAG 1074
DB 19957 GTGCTGATGAGCTGTGTGCTGCTGTGTGCTGATGATGTTGGTTAGTCCCGGACGAG 20016
QY 1075 GCAACCTGTGCTAATATGCGATCAATTTGTTGGGCACTTATGAGACTGCGCGGTGAC 1134
DB 20017 GCAACCTGTGCTAATATGCGATCAATTTGTTGGGCACTTATGAGACTGCGCGGTGAC 20076
QY 1135 AAACCGGAGAGGTTGGGATGACGTCAGTCCCTCATGAGCCCTTATGGGTTAGGGCTTAC 1194
DB 20077 AAACCGGAGAGGTTGGGATGACGTCAGTCCCTCATGAGCCCTTATGGGTTAGGGCTTAC 20136
QY 1195 AGCTAATACATGAGCGGTCAGAGAGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGA 1254
DB 20137 AGCTAATACATGAGCGGTCAGAGAGGTTGCCAACCCGCGAGGGGAGCTTAATCTCAGA 20196
QY 1255 AGGCGCTGCTAGTCCGAGTGGAGTCTGCACTCGCTGGAAGTCCGGAATCGCTAGT 1314
DB 20197 AGGCGCTGCTAGTCCGAGTGGAGTCTGCACTCGCTGGAAGTCCGGAATCGCTAGT 20256
QY 1315 AATCGCGATCAGCATGTCGCGGTGAATACGTTCCCGGCTTGTATACACACCGCGCTCA 1374
DB 20257 AATCGCGATCAGCATGTCGCGGTGAATACGTTCCCGGCTTGTATACACACCGCGCTCA 20316
QY 1375 CACCATGAGAGTGGGTTTACACAGAGAGGTAGTCTAAGCTTAAGAGAGGGCGTTTCCA 1434
DB 20317 CACCATGAGAGTGGGTTTACACAGAGAGGTAGTCTAAGCTTAAGAGAGGGCGTTTCCA 20376
QY 1435 CGGTGAGATTCATGACTGGGGTG 1457
DB 20377 CGGTGAGATTCATGACTGGGGTG 20399
```

```
RESULT 52
US-10-168-337A-5
; Sequence 5, Application US/10168337A
; Publication No. US20030170654A1
; GENERAL INFORMATION:
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholte, Philip; and Blackall,
; APPLICANT: Linda L.
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyporphosphate
; FILE REFERENCE: 002367PC/KF
; CURRENT APPLICATION NUMBER: US/10/168,337A
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1322
; TYPE: DNA
; ORGANISM: Rhodocyclus sp.
US-10-168-337A-5

Query Match 74.8%; Score 1089.4; DB 6; Length 1322;
Best Local Similarity 90.1%; Pred. No. 3.1e-280;
Matches 1190; Conservative 0; Mismatches 126; Indels 5; Gaps 2;

QY 142 TACCCGATATATCTTAAGAGAGGAGGAGGATCGAAAGACCTTGCGCTTTGAGCGG 201
DB 1 TACCCGATATATCTTAAGAGAGGAGGAGGATCGAAAGACCTTGCGCTTTGAGCGG 60
QY 202 CCGATGTCGATTAAGTATGTTGGTGGGTTAAAGGCTTACCAAGCGGACGATGATTTG 261
DB 61 CCGATGTCGATTAAGTATGTTGGTGGGTTAAAGGCTTACCAAGCGGACGATGATTTG 120
QY 262 GTCTGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 321
DB 121 GTCTGAGAGAGATGATCCGCGCACATGCGGACTGAGACACGCGCCAGACTCTTACGGGAGGC 180
QY 322 AGCAGTGGGGAATTTTGGCAATGGGCGCAACCTGATCAGCAATGCGCGTGAAGTGA 381
DB 181 AGCAGTGGGGAATTTTGGCAATGGGCGCAACCTGATCAGCAATGCGCGTGAAGTGA 240
QY 382 GAAGGCTTGGGTTGTAAAGCTCTTTCAGTGAAGAAAGTTACGTTAATTAATCG 441
DB 241 GAAGGCTTGGGTTGTAAAGCTCTTTCAGTGAAGAAAGTTACGTTAATTAATCG 300
QY 442 TGAATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501
DB 301 TGAATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
QY 502 ATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
DB 361 ATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 562 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621
DB 421 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 622 AGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 681
DB 481 AGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
QY 682 AGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 741
DB 541 AGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 742 GGGAGCAAAACAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 801
DB 601 GGGAGCAAAACAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 802 TTGGGCTTATTAAGCTTG--GTAAAGAACTTAAGCTTAAGCTTAAGCTTAAGCTTAAG 859
DB 661 TTGGGAGGTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAA 720
QY 860 ACGGTGCAAGATTAATAAATCAAAAGAAATTGAAGGGGAGCCCGCAAGCGGTGATTAATG 919
```

Db 721 ACGGCGGAGGCTAAAAGTCAAGAAATGACGGGAGCCGCAACAAGGGTGAATG 780  
Qy 920 TGAATTAATTCGATGCAACGCGAAGAACTTAACCTTACCTTGAATGAGCAATTTCT 979  
Db 781 TGAATTAATTCGATGCAACGCGAAGAACTTAACCTTGAATGAGCAATTTCT 840  
Qy 980 AGAGATTAATTAATTCGATGCAACGCGAAGAACTTAACCTTGAATGAGCAATTTCT 1036  
Db 1037 TCGTGTGATGAGATGTTGGGTTAAGTCCGCAACGAGCGCAACCTTGTCAATTAATG 1096  
Qy 901 TCGTGTGATGAGATGTTGGGTTAAGTCCGCAACGAGCGCAACCTTGTCAATTAATG 960  
Db 1097 ATCATTTGGTGGGCACTTTAATGAGACTGCGGTGACAAACCGAGAGAGTGGGATG 1156  
Qy 961 ATCATTTGGTGGGCACTTTAATGAGACTGCGGTGACAAACCGAGAGAGTGGGATG 1020  
Db 1157 ACGTCAATGCTCATGAGGCTTATGAGGCTTCAACGATTAATCAATGGCGGTACA 1216  
Qy 1021 ACGTCAATGCTCATGAGGCTTATGAGGCTTCAACGATTAATCAATGGCGGTACA 1080  
Db 1217 GAGGTTGCGCAACCGCGAGGAGGAGCTAATCTCAAGAAACCGGTCTGATCGGATCG 1276  
Qy 1081 GAGGTTGCGCAACCGCGAGGAGGAGCTAATCTCAAGAAACCGGTCTGATCGGATCG 1140  
Db 1277 AGTCTGCACTGACTCCGTGAATGCGAATGCTAATGATGCGGATCAGATCGCG 1336  
Qy 1141 AGTCTGCACTGACTCCGTGAATGCGAATGCTAATGATGCGGATCAGATCGCG 1200  
Db 1337 GTGAATAGGTTCCGGGCTTGTACACACCGCGCTCACACCATGAGAGGAGTTCAC 1396  
Qy 1201 GTGAATAGGTTCCGGGCTTGTACACACCGCGCTCACACCATGAGAGGAGTTCAC 1260  
Db 1397 AGAAGCAGATGCTTAACGTAAGAGAGGCGCTTGCACGAGATTCATGATCGGAGT 1456  
Qy 1261 AGAAGCAGATGCTTAACGTAAGAGAGGCGCTTGCACGAGATTCATGATCGGAGT 1320  
Qy 1457 G 1457  
Db 1321 G 1321

RESULT 53  
US-10-464-724-e/c  
; Sequence 6, Application US/10464724  
; Publication No. US20030203398A1  
; GENERAL INFORMATION:  
; APPLICANT: Brummett, Michael  
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
; FILE REFERENCE: BC1033 US NA  
; CURRENT APPLICATION NUMBER: US/10/464,724  
; PRIOR FILING DATE: 2003-06-17  
; PRIOR APPLICATION NUMBER: US/09/735,567  
; PRIOR FILING DATE: 2000-12-13  
; PRIOR APPLICATION NUMBER: 60/171,140  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 1539  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Brachymonas  
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RA6  
US-10-464-724-6

Query Match 74.2%; Score 1080.8; DB 6; Length 1539;  
Best Local Similarity 85.7%; Pred. No. 6,4e-276;

Matches 1251; Conservative 0; Mismatches 197; Indels 12; Gaps 4;  
Qy 3 TGAAGCTGGGCGGATGCTTAAACATGCAAGTGAAGCGGAGACGGATGCTGCAATCT 62  
Db 1538 TGAAGCTGGGCGGATGCTTAAACATGCAAGTGAAGCGGAGACGGATGCTGCAATCT 1481  
Qy 63 GGTGGCGAGTGGCGGACGGGTGAGTAATGATTCGGAACGTAATTCAGAAAGAGGGGGTAAC 122  
Db 1480 GGTGGCGAGTGGCGGACGGGTGAGTAATGATTCGGAACGTAATTCAGAAAGAGGGGGTAAC 1421  
Qy 123 GCATCGAAAGATGTCCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATC--GAAA 180  
Db 1420 TCGGCGAAAGCGGATTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATC--GAAA 1361  
Qy 181 GACCTGCGCTTTTGAAGCGGCGGATGCTGATTAAGTATGAGTGGGCTTAAAGGCTTAC 240  
Db 1360 GACCTGCGCTTTTGAAGCGGCGGATGCTGATTAAGTATGAGTGGGCTTAAAGGCTTAC 1301  
Qy 241 CAAGCGACGATCAGTATGCTGAGAGAGACGACGACCACTGGGACTGAGACAG 300  
Db 1300 CAAGCGACGATCAGTATGCTGAGAGAGATGATACGACCACTGGGACTGAGACAG 1241  
Qy 301 GCCCAACTCTTACGAGAGGAGAGAGAGATTTTGAACATGAGGCGCAAGCTGATC 360  
Db 1240 GCCCAACTCTTACGAGAGGAGAGAGAGATTTTGAACATGAGGCGCAAGCTGATC 1181  
Qy 361 CAGCAATGCGCGGTGAGTAAGAGAGGCTTGGGCTTAAAGCTCTTCACTGAGAAAGA 420  
Db 1180 CAGCAATGCGCGGTGAGTAAGAGAGGCTTGGGCTTAAAGCTCTTCACTGAGAAAGA 1121  
Qy 421 AAAGCTTACGCTTAATTAATCTGATCTGATGACGATGACAGAAAGACACGCGCTAAC 480  
Db 1120 AAAGCTTCTTCTAATTAAGAGGACATGACGATGACGATTAAGAAATTAAGACCGGCTAAC 1061  
Qy 481 TACGTCCAGACGCGCGGTAAATACGTAAGGTGCAAGCGTTAATCGGAATTAATCTGCGCT 540  
Db 1060 TACGTCCAGACGCGCGGTAAATACGTAAGGTGCAAGCGTTAATCGGAATTAATCTGCGCT 1001  
Qy 541 AAAGGTGCGGAGCGGCTTGTAAAGTCAATGTAATGTAATGCGGCGCTTAACTGGGAAT 600  
Db 1000 AAAGGTGCGGAGCGGCTTGTAAAGTCAATGTAATGTAATGCGGCGCTTAACTGGGAAT 941  
Qy 601 TCGTTTGAACCTAACAGGCTAGAGTGTGCGAGAGGAGTGAATTCATGTTAGCAG 660  
Db 940 TCGTTTGAACCTAACAGGCTAGAGTGTGCGAGAGGAGTGAATTCATGTTAGCAG 881  
Qy 661 TGAATGCTGATGATGAGTAAGAAACATGATGCGAAAGGAGCGCTCTGGGTTAACT 720  
Db 880 TGAATGCTGATGATGAGTAAGAAACATGATGCGAAAGGAGCGCTCTGGGTTAACT 821  
Qy 721 GACGCTCATGACGAAAGCGGAGGAGCAAGATTAAGTACCTGCTGATGACGCGC 780  
Db 820 GACGCTCATGACGAAAGCGGAGGAGCAAGATTAAGTACCTGCTGATGACGCGC 761  
Qy 781 CTAACGATGCTCACTAGTGTGGGCTTAAAGCTTGAAGCTGATTAAGCTAACGCGTGA 840  
Db 840 CTAACGATGCTCACTAGTGTGGGCTTAAAGCTTGAAGCTGATTAAGCTAACGCGTGA 701  
Qy 841 AGTTGACCGCTGGGAGATGATGCGCAAGTAACTCAAGAGATTAAGCGGAGACCC 900  
Db 700 AGTTGACCGCTGGGAGATGATGCGCGCAAGTAACTCAAGAGATTAAGCGGAGACCC 641  
Qy 901 GCACAAGCGGTGATTAATGAGTAAATGATGATGCAACGAGAAACCTTACCTTACCCTT 960  
Db 640 GCACAAGCGGTGATTAATGAGTAAATGATGATGCAACGAGAAACCTTACCTTACCCTT 581  
Qy 961 GACATGAGCGAATTTCTAAGATTAATTAATGCT--TCGGAGCGCTTAACACAGGTG 1017  
Db 580 GACATGAGCGAATTTCTAAGATTAATTAATGCT--TCGGAGCGCTTAACACAGGTG 521  
Qy 1018 CTGATGCTGTGCTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1077  
Db 520 CTGATGCTGTGCTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 461



```
Db      285 CGGTGATGTCGGATTCGCAAGTCTGCACTGACGCGGTAAGTCGGAATGCTAGTAAT 226
      |||
Qy      1318 GCGGATCAAGCATGTTCGGGTAATACGTTCCCGGGCTCTTTCACACACCGCCCTGACAC 1377
      |||
Db      225 CGTGATCAGCATGTTCACGGTGAATACGTTCCCGGGCTCTTTCACACACCGCCCTGACAC 166
      |||
Qy      1378 CATGGAGTGGGTTTACCAAGAACGAGTGTCTAACCGTAAAGAGGGCGCTTGCACAG 1437
      |||
Db      165 CATGGAGAGGGGTTCTCACCAAGAACGAGTGTCTAACCGCAAGAGGGCGCTTGCACAG 106
      |||
Qy      1438 TGAGATTCATGACTGGGGTG 1457
      |||
Db      105 TGGGGTTCTGACTGGGGTG 86
      |||

RESULT 55
US-10-464-709-6/c
; Sequence 6, Application US/10464709
; Publication No. US20030207321A1
; GENERAL INFORMATION:
; APPLICANT: Brameccl, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE REFERENCE: BC1033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,709
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 1539
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Brachyomonas
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RA6
US-10-464-709-6

Query Match      74.2% Score 1080.8; DB 6; Length 1539;
Beet Local Similarity 85.7%; Pred. No. 6.4e-278;
Matches 1251; Conservative 0; Mismatches 197; Indels 12; Gaps 4;

Qy      3 TGAACGCTGCGCGCATGCTTTACACATGCAAGTCGACGCGGACGACGAGTGGTTCATCT 62
      |||
Db      1538 TGAACGCTGCGCGCATGCTTTACACATGCAAGTCGACGCGGACGCG--TCCTTGGGAT 1481
      |||
Qy      63 GGTGCGAGTGGCGGACGCGGTGATGATGATCGAAGCGATTCGAAGAGGGGGTAAAC 122
      |||
Db      1480 GCGCGCGAGTGGCGAAGCGGGTGAATGAAGCATCGAAGCGTCCGCTGATGGGATAGC 1421
      |||
Qy      123 GCATGGAAGATGTCTAATACCGCATATACCTTAAGAGAGAAACAGGGGATC--GAAA 180
      |||
Db      1420 TCGGGGAAGCCGGAATTAATACCGCATAGATCCGTGATGAAGCAAGGGGACCCGCAAG 1361
      |||
Qy      181 GACCTTGGCTTTTGGAGCGGCGCATGCTGTGATTAAGTGTGGGGTAAAGGCTTAC 240
      |||
Db      1360 GGCCTTGGCTTACTGAGACGCGCATGTCAATTAAGTGTGGGGTAAAGGCGCAC 1301
      |||
Qy      241 CAAGCGACGATCACTAGTGTGTGAGAGACGACCACTCTGGGACTGAGACAG 300
      |||
Db      1300 CAAGCTCGATCTGTAGCTGGTGAAGAGATGATCAAGCCACTCGGAGCTGAAGACAG 1241
      |||
Qy      301 GCCCGACTCTTACGCGAGCGGACGATGGGAAATTTTGAACAATGGGCGCAAGCCTGATC 360
      |||
Db      1240 GCCCGACTCTTACGCGAGCGGACGATGGGAAATTTTGAACAATGGGCGCAAGCCTGATC 1181
      |||
Qy      361 CAGCAATCCCGCTGATGAAGAAAGGCTTGGGGTGTAAAGCTTTCACTGAGAAAGA 420
      |||
```

```
Db      1180 CAGCCATGCCCGCTGACAGATGAAGGCTTCGGGTTGTAACTGCTTTTGTACGAAACGA 1121
      |||
Qy      421 AAGGTTACGGTAAATATATGTCATGACGGTATGACAGAAAGACACCGGCTTAC 480
      |||
Db      1120 AAGGCTCTTTCTTAATTAAGGGCACTGACGGTTACCGTAAGAAATTAAGCACCGGCTTAC 1061
      |||
Qy      481 TACGTGCCAGCAGCGCGGTAATACGTAGGGTGCAGCGTTAATCGAAATTACTGGGCGT 540
      |||
Db      1060 TACGTGCCAGCAGCGCGGTAATACGTAGGGTGCAGCGTTAATCGAAATTACTGGGCGT 1001
      |||
Qy      541 AAGGGTGCAGCGCGGCTTTGTATGATGATGTGAATCCCGGGCTTAATCTGGGAAT 600
      |||
Db      1000 AAGCGTGCAGCGCGGCTTTGTGACAGTGTGAATCCCGGGCTTAATCTGGGAATC 941
      |||
Qy      601 TGGCTTTGAATCTAAGGCTAGAGTGTGCAGAGGAGGATGGAATTCATGTGTAGAG 660
      |||
Db      940 TGGCATTTGTGACTCAAGGCTGAGTACGGCAGAGGAGGAGATGGAATTCGCGTGTAGAG 881
      |||
Qy      661 TGAATGCGTGAATATGGAAGAACATCGATGCGAAGGACGCTCTGGGTTAAGACT 720
      |||
Db      880 TGAATGCGTGAATATGCGAGAGAACACCGATGCGAAGGACGCTCTGGGCTGTACT 821
      |||
Qy      721 GACGCTATGACGAAAGCGTGGGAGCAAAACAGATTAATACCTGTGATGTCACGCC 780
      |||
Db      820 GACGCTATGACGAAAGCGTGGGAGCAAAACAGATTAATACCTGTGATGTCACGCC 761
      |||
Qy      781 CTAACGATGTCACTGTTGTGGGCTTAATTAAGGCTGTGTAAGGCTTAACGCGTGA 840
      |||
Db      760 CTAACGATGTCACTGTTGTGGGCTTAATTAAGGCTGTGTAAGGCTTAACGCGTGA 701
      |||
Qy      841 AGTTGACCGCTGGGAGTACGATCGCAAGATTAATACTCAAGAAATGTAACGGGACCC 900
      |||
Db      700 AGTTGACCGCTGGGAGTACGATCGCAAGATTAATACTCAAGAAATGTAACGGGACCC 641
      |||
Qy      901 GCACAGCGGTGATTAATGTGATTAATGATGACAAACGGAATAACCTTAACCTT 960
      |||
Db      640 GCACAGCGGTGATTAATGTGATTAATGATGACAAACGGAATAACCTTAACCTT 581
      |||
Qy      961 GACATGTAGCGAATTTTCTAGAGTAAATTAAGTCTCTGAGGAAAGCTTAACACAGGTG 1017
      |||
Db      580 GACATGTAGCGAATTTTCTAGAGTAAATTAAGTCTCTGAGGAAAGCTTAACACAGGTG 521
      |||
Qy      1018 CTGATGAGCTGTCTGACGCTGTGTCGTGAGATGTTGGTAAATCCCGCAACGAGCGCA 1077
      |||
Db      520 CTGATGAGCTGTCTGACGCTGTGTCGTGAGATGTTGGTAAATCCCGCAACGAGCGCA 461
      |||
Qy      1078 ACCCTTGTCAATTAATTCGCATCAATTTGGTGGCACTTTAATGATGATCGCGGTGACAA 1137
      |||
Db      460 ACCCTTGTCAATTAATTCGCATCAATTTGGTGGCACTTTAATGATGATCGCGGTGACAA 406
      |||
Qy      1138 CCGGAGGAAGGTGGGATGACGTCAATGCTCAATGCGCTTAATGGGTAAGGCGTTACACAG 1197
      |||
Db      405 CCGGAGGAAGGTGGGATGACGTCAATGCTCAATGCGCTTAATGGGTAAGGCGTTACACAG 346
      |||
Qy      1198 TAATACATGGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTTAATCTGAAAGC 1257
      |||
Db      345 TCAATCAATGGCGGTACAGAGGTTGCAACCCGCGAGGGGAGCTTAATCTGAAAGC 286
      |||
Qy      1258 GCGTGTAGTCCGAGTCGAGTGTGCAATCTGCACTCGGTGAAGTGGAAATGCTTAAT 1317
      |||
Db      285 GCGTGTAGTCCGAGTCGAGTGTGCAATCTGCACTCGGTGAAGTGGAAATGCTTAAT 226
      |||
Qy      1318 GCGGATCAAGATGTCCGCGTGAATTAAGTTCCGCGGCTTGTACACACCGCCCTGACAC 1377
      |||
Db      225 CGTGATCAAGATGTCCGCGTGAATTAAGTTCCGCGGCTTGTACACACCGCCCTGACAC 166
      |||
Qy      1378 CATGGAGTGGGTTTCAACGAAGACAGTGTCTTAACGTAAGAGGGGCGCTTGCACAG 1437
      |||
Db      165 CATGGAGCGGGTTCTCACGAAGACAGTGTCTTAACGTAAGAGGGGCGCTTGCACAG 106
      |||
Qy      1438 TGAGATTCATGACTGGGGTG 1457
      |||
Db      105 TGGGGTTCTGACTGGGGTG 86
      |||
```

```
RESULT 56
US-10-464-724-7
; Sequence 7, Application US/10464724
; Publication No. US20030203398A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; TITLE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BC1033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,724
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Brachymonas
; OTHER INFORMATION: Nucleotide sequence of the 16S rDNA corresponding to strain RA9
US-10-464-724-7

Query Match      74.1%; Score 1080.2; DB 6; Length 1454;
Best Local Similarity 85.6%; Pred. No. 9.1e-278;
Matches 1251; Conservative 0; Mismatches 198; Indels 12; Gaps 4;

QY      2 TTGAAGCTTGGCGGCGATGCTTTACATGCAAGTCCGAAGCGGAGCGATGCTTGATC 61
DB      1 TTGAAGCTTGGCGGCGATGCTTTACATGCAAGTCCGAAGCGGAGCGGATGCTTGATC 58
QY      62 TGGTGGCGAGTGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121
DB      59 TGGTGGCGAGTGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 118
QY      122 CGCATGAAAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 179
DB      119 CTGGGCGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 178
QY      180 AGACCTTGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 239
DB      179 GGGCTTGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 238
QY      240 CCAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299
DB      239 CCAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298
QY      300 GGCCGAGATCTTCAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359
DB      299 GGCCGAGATCTTCAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 358
QY      360 CCAGCAATGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419
DB      359 CCAGCAATGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418
QY      420 AAAAGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
DB      419 AAAAGCTTCTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 478
QY      480 CTAGTGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
DB      479 CTAGTGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538
QY      540 TAAAGGCTGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
DB      539 TAAAGGCTGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 598
```

```
QY      600 TTGCGTTGAAACTCAAGGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
DB      599 CTGCGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 658
QY      660 GTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
DB      659 GTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
QY      720 TGAAGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
DB      719 TGAAGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778
QY      780 CCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
DB      779 CCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838
QY      840 AAGTTACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
DB      839 AAGTTACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 898
QY      900 CGCAGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
DB      899 CGCAGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
QY      960 TGAAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016
DB      959 TGAAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018
QY      1017 GCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
DB      1019 GCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078
QY      1077 AACCTTGCATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136
DB      1079 AACCTTGCATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133
QY      1137 ACCGAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196
DB      1134 ACCGAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
QY      1197 GTAATCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256
DB      1194 GTAATCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
QY      1257 CCGGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1316
DB      1254 CCGGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
QY      1317 TCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1376
DB      1314 TCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373
QY      1377 CCAATGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1436
DB      1374 CCAATGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1433
QY      1437 GTGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1497
DB      1434 GTGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1494
```

```
RESULT 57
US-10-464-356-7
; Sequence 7, Application US/10464356
; Publication No. US20030207320A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; TITLE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BC1033 US NA
; CURRENT APPLICATION NUMBER: US/10/464,356
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
```

;; PRIOR FILING DATE: 2000-12-13  
;; PRIOR APPLICATION NUMBER: 60/171,140  
;; PRIOR FILING DATE: 16 DECEMBER 1999  
;; NUMBER OF SEQ ID NOS: 30  
;; SOFTWARE: Microsoft Office 97  
;; SEQ ID NO 7  
;; LENGTH: 1454  
;; TYPE: DNA  
;; ORGANISM: Unknown Organism  
;; FEATURE: Description of Unknown Organism: Brachymonas  
;; OTHER INFORMATION: Nucleotide sequence of the 16S rDNA corresponding to strain RA9  
US-10-464-356-7

Query Match 74.1%; Score 1080.2; DB 6; Length 1454;  
Best Local Similarity 85.6%; Pred. No. 9.1e-278;  
Matches 1251; Conservative 0; Mismatches 198; Indels 12; Gaps 4;

QY 2 TTGAACGGCTGGCGCATGCTTACACATGCAAGTGGACGCGACCGGATGCTTGACATC 61  
DB 1 TTGAACGGCTGGCGCATGCTTACACATGCAAGTGGACGCGACCGGATGCTTGACATC 58  
QY 62 TGGTGGCGAGTGGCGGACGCGGTGAGTATGATCGGAACGTATCCAGAAAGAGGGGGTAA 121  
DB 59 TGGCGGCGAGTGGCGGACGCGGTGAGTATGATCGGAACGTATCCCGGTATGTCGGGATAG 118  
QY 122 CGCATCGAAAGATGTGCTATACCGCATATCTTAAAGAGAAACGAGGGA--TCGAA 179  
DB 119 CTCGGCGAAAGCGGATTAATACCGCATGATCCGTGATAAAGCAAGGGGACTCGCA 178  
QY 180 AGACCTTGGCTTTTGGAGCGGCGCATGCTGATTAAGTGGTGGGGTAAAGGCTTA 239  
DB 179 GGGGCTTGGCTTATGAGCGGCGCATGCTGATTAAGTGGTGGGGTAAAGGCTTA 238  
QY 240 CCAAGGCGACGATCAGTATGCTGAGAGACGACGACCACTGGGACTGAGAC 299  
DB 239 CCAAGGCGACGATCAGTATGCTGAGAGACGACGACCACTGGGACTGAGAC 298  
QY 300 GGGCCGACCTCTTACGCGGAGGACGAGCATGCGGGAATTTTGAACATGCGCGACGCTGAT 359  
DB 299 GGGCCGACCTCTTACGCGGAGGACGAGCATGCGGGAATTTTGAACATGCGCGACGCTGAT 358  
QY 360 CCAGCAATGCGCGGAGTGAAGAAAGGCTTGGGTTGAAAGCTTTCACTGAGAG 419  
DB 359 CCAGCAATGCGCGGAGTGAAGAAAGGCTTGGGTTGAAAGCTTTCACTGAGAG 418  
QY 420 AAAAGGTAACGCTAATATCGTACATGACGCTATCGACAGAAAGCACCGGCTTA 479  
DB 419 AAAAGGCTCTTCTAATTAAGAGGACATGACGCTACCGTAAGAAATTAAGCACCGGCTTA 478  
QY 480 CTACGTCGACGACGCGCGGTAAATCGTAGGGTGAAGGCTTAATCGAATTAATCTGGGCG 539  
DB 479 CTACGTCGACGACGCGCGGTAAATCGTAGGGTGAAGGCTTAATCGAATTAATCTGGGCG 538  
QY 540 TAAAGGCGGCGGCGGCTTTGTAAGTGAAGTGAAGTCCCGGGCTTAACCTGGGAA 599  
DB 539 TAAAGGCGGCGGCGGCTTTGTAAGTGAAGTGAAGTCCCGGGCTTAACCTGGGAA 598  
QY 600 TTGCGTTTGAACCTCAAGGCTGAGTGGCAGAGGAGGTGAATTCATGTTAGACA 659  
DB 599 CTGCGATTGTGACTGCAAGGCTGAGTACGCGAGAGGGGATGAATTCGCGGTAGACA 658  
QY 660 GTGAATCGTAGAGATATGAAAGAAATCGATGCGGAAAGGCGCTCTCGGTTAAAC 719  
DB 659 GTGAATCGTAGAGATATCGGAGAAACACCGATGCGGAAAGGCGCTCTCGGTTAAAC 718  
QY 720 TGACGCTATGACGAAAGCGTGGGAGCAACAGAGTTAATACCTGTTAGTCCAGC 779  
DB 719 TGACGCTATGACGAAAGCGTGGGAGCAACAGAGTTAATACCTGTTAGTCCAGC 778  
QY 780 CCTAAACGATGCACTAGTTGTTGGGCTTATTAAGGCTTGGTAAAGCAAGCTAACGCGTG 839

DB 779 CCTAAACGATGCACTGTTGTTGGGATTTATTTCTTCAGTAACGAACTTAACGCGTG 838  
QY 840 AAGTTACCGGCTGGGAGTACGGTCCGCAAGATTAACTCAAGAAATTGACGGGAC 899  
DB 839 AAGTTACCGGCTGGGAGTACGGTCCGCAAGATTAACTCAAGAAATTGACGGGAC 898  
QY 900 CGCACAAGCGGTGATTTATGATTAATTTGATGCAACGCGAAACCTTACTACCT 959  
DB 899 CGCACAAGCGGTGATTTATGATTTATTTGATGCAACGCGAAACCTTACTACCT 958  
QY 960 TGACATGAGCGAAATTTCTAGAGATGATTAAGTCT--TCGGAACGCTAACAGGT 1016  
DB 959 TGACATGATGAATATCCCGCAGAGATGGGAGTCTCGCAAGAGAGCCATTAACAGGT 1018  
QY 1017 GCTGATGCTGTGTGACGCTCGTGTGTGATGATTTAGGTTAGTCCCGCAAGAGCGC 1076  
DB 1019 GCTGATGCTGTGTGACGCTCGTGTGTGATGATTTAGGTTAGTCCCGCAAGAGCGC 1078  
QY 1077 AACCTTGTCAATTAATTCATCTATTTGTTGGGCACTTTAATGACTGCGGTTGACAA 1136  
DB 1079 AACCTTGTCAATCTAGTTGCTACGAAA-----GGGCACTGTATGGGACTGCGGTTGACAA 1133  
QY 1137 ACCGAGGAGTGGGAGTGAAGTCAAGTCTCATGTCGCTTATGAGGCTTACAC 1196  
DB 1134 ACCGAGGAGTGGGAGTGAAGTCAAGTCTCATGTCGCTTATGAGGCTTACAC 1193  
QY 1197 GTAATTAAGTGGCGGTACAGAGGTTGCCAACCGCGGAGGAGGCTTAATTCAGAAAG 1256  
DB 1194 GTCAATCAATGCGCGGTACAGAGGTTGCCAACCGCGGAGGAGGCTTAATTCAGAAAG 1253  
QY 1257 CGGCTGTGATGCGGATGCGAGTCTGCAACTGCACTCGTGAATCGGAAATCGCTAGTAA 1316  
DB 1254 CGGCTGTGATGCGGATGCGAGTCTGCAACTGCACTCGTGAATCGGAAATCGCTAGTAA 1313  
QY 1317 TCGCGATCAGCATGTCGCGGTAAATGTTCCGCGGCTTTTGAACACCGCCGTCACA 1376  
DB 1314 TCGGTATCAGCATGTCAGGTGAATGTTCCGCGGCTTTTGAACACCGCCGTCACA 1373  
QY 1377 CCATGGAGTGGGTTTCAACGAAAGCAGTATGTTAAACGTAAGAGGCGCTTGGCAGC 1436  
DB 1374 CCATGGAGTGGGTTTCAACGAAAGCAGTATGTTAAACGTAAGAGGCGCTTGGCAGC 1433  
QY 1437 GTGAGATTCAATGACTGGGGTG 1457  
DB 1434 GTGGGTTTCTGACTGGGGTG 1454

RESULT 58  
US-10-464-709-7  
; Sequence 7, Application US/10464709  
; Publication No. US20030207321A1  
; GENERAL INFORMATION:  
; APPLICANT: Branucci, Michael  
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
; FILE REFERENCE: BCI003 US NA  
; CURRENT APPLICATION NUMBER: US/10/464,709  
; CURRENT FILING DATE: 2003-06-17  
; PRIOR APPLICATION NUMBER: US/09/735,567  
; PRIOR FILING DATE: 2000-12-13  
; PRIOR APPLICATION NUMBER: 60/171,140  
; PRIOR FILING DATE: 16 DECEMBER 1999  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 1454  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE: Description of Unknown Organism: Brachymonas  
; OTHER INFORMATION: Nucleotide sequence of the 16S rDNA corresponding to strain RA9  
US-10-464-709-7

Query Match 74.1%; Score 1080.2; DB 6; Length 1454;  
Best Local Similarity 85.6%; Pred. No. 9.1e-278;  
Matches 1251; Conservative 0; Mismatches 198; Indels 12; Gaps 4;

QY 2 TTGAACGCTGGCGGCGATGCTTTACACATTCGAAGTCGACGCGACGAGTCCTGCATC 61  
DB 1 TTGAACGCTGGCGGCGATGCTTTACACATTCGAAGTCGACGCGACGCGCG--TCCTTCGGGA 58  
QY 62 TGGTGGCGGATGGCGGACGGGTGAGTAATGATCGGAAGCTATCCGAAGAGGGGGTAA 121  
DB 59 TGGCGCGAGTGGCGGACGGGTGAGTAATGATCGGAAGCTATCCGAAGAGGGGGTAA 118  
QY 122 CGCATGGAAGAATGTCCTAATACCGCATATCTTAAGAGGAAACGAGGGGA--TCGA 179  
DB 119 CTCGGCGAAGCCGGAATTAAATCCGATGATCCGTGATGAAGCAAGGGGATCGCA 178  
QY 180 AGACCTTGGCTTTTGGAGCGGCGGATGCTGATTAAGTATGTTGGTGGGTAAAGGCTTA 239  
DB 179 GGGCTTGGCTTACTGAGCGGCGGATGCTGATTAAGTATGTTGGTGGGTAAAGGCTTA 238  
QY 240 CCAAGCGGACGATCACTAGTGGTCTGAGAGGACGACGCGACATCGGGACTGAGACAC 299  
DB 239 CCAAGCTGCGATCTGATCTGATGAGTGAATGATGACCATCGGGACTGAGACAC 298  
QY 300 GGGCGGACCTCTTAACGGAGGACGAGTGGGGAATTTTGGACAAATGGGCGCAAGCTGAT 359  
DB 299 GGGCGGACCTCTTAACGGAGGACGAGTGGGGAATTTTGGACAAATGGGCGCAAGCTGAT 358  
QY 360 CCAGCAATGCGCGGTGATGAAGAGGCGCTTGGGTTGTAAGCTCTTTCATCGAAGAG 419  
DB 359 CCAGCAATGCGCGGTGATGAAGAGGCGCTTGGGTTGTAAGCTCTTTCATCGAAGAG 418  
QY 420 AAAAGCTTACGTAATTAATCGTCACTCAATGAGGTATGACGAAGAGGACGCGCTAA 479  
DB 419 AAAAGCTTCTTCTAATTAAGAGGACCATGACGATGATGTAAGATTAAGCACCGCTAA 478  
QY 480 CTACGTGCGACGACGCGCGGTAAATGATGAGGTGCAAGCTTATCGAATTAATCGGCG 539  
DB 479 CTACGTGCGACGACGCGCGGTAAATGATGAGGTGCAAGCTTATCGAATTAATCGGCG 538  
QY 540 TAAAGGTGCGAGCGCGCTTGTAAAGTCAATGTGAATCCCGGCTTAACTTGGGAA 599  
DB 539 TAAAGGTGCGAGCGCGCTTGTAAAGTCAATGTGAATCCCGGCTTAACTTGGGAA 598  
QY 600 TTGCGTTTGAATCAAGGCTAGAGTGGGACGAGGAGGAGTGAATTCATGATGATGCA 659  
DB 599 CTGCGCAATGTGACTGCAAGGCTGAGATGCGCAGAGGGGATGGAATTCGCGGTGACA 658  
QY 660 GTGAATGCGTAGATGATGAAGAAATCATGATGCGCAAGGCGCTCTGCGTTAACAC 719  
DB 659 GTGAATGCGTAGATGATGAAGAAATCATGATGCGCAAGGCGCTCTGCGTTAACAC 718  
QY 720 TGAAGCTCATGACGAAGAGCGTGGGAGCAAAACAGATTAGATCCCTGTAGTCCAGC 779  
DB 719 TGAAGCTCATGACGAAGAGCGTGGGAGCAAAACAGATTAGATCCCTGTAGTCCAGC 778  
QY 780 CTTAAACGATGATCACTAGTGTGGGCTTATTAAGCTTGTGATGCAAGCAATCGGCTG 839  
DB 779 CTTAAACGATGATCACTAGTGTGGGCTTATTAAGCTTGTGATGCAAGCAATCGGCTG 838  
QY 840 AAGTTGACCGCTGGGAGTACCGGTCCGAGATTAATAACTCAAGGAATTTGACGGGAC 899  
DB 839 AAGTTGACCGCTGGGAGTACCGGTCCGAGATTAATAACTCAAGGAATTTGACGGGAC 898  
QY 900 CGCAAGAGCGGTGATTAATGATTAATTCGATGCAACGCGAATACTTATCTTACCT 959  
DB 899 CGCAAGAGCGGTGATTAATGATTAATTCGATGCAACGCGAATACTTATCTTACCT 958  
QY 960 TGACATGATGAGAAATTTTCTAGAGATGATTAAGTCT---TCGGGAACGCTTAACAGCT 1016  
DB 959 TGACATGATGAGAAATTTTCTAGAGATGATTAAGTCT---TCGGGAACGCTTAACAGCT 1018

QY 1017 GCTGATGCGCTGTGTCAGCTCGTGTGAGATGTTGGTTAAAGTCCCGCAACGAGCGC 1076  
DB 1019 GCTGATGCGCTGTGTCAGCTCGTGTGAGATGTTGGTTAAAGTCCCGCAACGAGCGC 1078  
QY 1077 AACCTTGTCAATTAATTCATCATTTTGGGACCTTATAGACTGCGGTGACAA 1136  
DB 1079 AACCTTGTCAATTAATTCATCATTTTGGGACCTTATAGACTGCGGTGACAA 1133  
QY 1137 ACCGAGAGAGTGGGATGACGTCAATCTCATAGCCCTTATAGGAGGCTTACAC 1196  
DB 1134 ACCGAGAGAGTGGGATGACGTCAATCTCATAGGCTTATAGGAGGCTTACAC 1193  
QY 1197 GTATATCAATGCGCGTACAGAGGTGCAACCCGAGAGGAGCTTAATCTCAAG 1256  
DB 1194 GTATATCAATGCGCGTACAGAGGTGCAACCCGAGAGGAGCTTAATCTCAAG 1253  
QY 1257 CGCGTGTAGTCCGATCGAGTCTGCACTCGATCCGTGAAGTCCGAATCGTATGA 1316  
DB 1254 CGCGTGTAGTCCGATCGAGTCTGCACTCGATCCGTGAAGTCCGAATCGTATGA 1313  
QY 1317 TCGCGATCAATGCTCGCGGTGAATTAAGTCCCGGCTTGTATACACCCGCGTACA 1376  
DB 1314 TCGCGATCAATGCTCGCGGTGAATTAAGTCCCGGCTTGTATACACCCGCGTACA 1373  
QY 1377 CCATGGAGTGGGTTTACCAAGACAGGTAGCTTAACGTAAGAGGCGCTTGCACG 1436  
DB 1374 CCATGGAGTGGGTTTACCAAGACAGGTAGCTTAACGTAAGAGGCGCTTGCACG 1433  
QY 1437 GTGATGATGATGAGGCTG 1457  
DB 1434 GTGATGATGATGAGGCTG 1454

RESULT 59  
US-10-464-724-2/c  
; Sequence 2, Application US/10464724  
; Publication No. US20030203398A1  
; GENERAL INFORMATION:  
; APPLICANT: Bramecl, Michael  
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
; TITLE OF INVENTION: Industrial Wastewater Bioreactors  
; FILE REFERENCE: BC1033 US NA  
; CURRENT APPLICATION NUMBER: US/10/464,724  
; PRIOR FILING DATE: 2003-06-17  
; PRIOR APPLICATION NUMBER: US/09/735,567  
; PRIOR FILING DATE: 2000-12-13  
; PRIOR APPLICATION NUMBER: 60/171,140  
; PRIOR FILING DATE: 16 DECEMBER 1999  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 1467  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Brachyomonas  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain NBYE13  
US-10-464-724-2

Query Match 74.1%; Score 1079.6; DB 6; Length 1467;  
Best Local Similarity 85.6%; Pred. No. 1.3e-277;  
Matches 1251; Conservative 0; Mismatches 199; Indels 12; Gaps 4;

QY 1 ATGAACGCTGGCGGCGATGCTTTACACATTCGAAGTCGACGCGACGAGTCCTGCATC 60  
DB 1465 ATGAACGCTGGCGGCGATGCTTTACACATTCGAAGTCGACGCGACGAGCGG--TCCTTCGGG 1408  
QY 61 CTGTTGCGAGTGGCGGACGGGTGAGTAATGATCGGAAGCTATCCGAAGAGGGGGTAA 120  
DB 1407 ATGCGCGGAGTGGCGGACGGGTGAGTAATGATCGGAAGCTATCCGAAGAGGGGGTAA 1348  
QY 121 ACGCATCAAGAAATGTCCTAATACCGCATATCTTAAGAGGAAAGCAGGGGA--TCGA 178



```

1347 GCTGCGCAAAACCGGATTTAAATCCGCATGAGATCCGTGAGTGAAGAAAGCAGGGGACCTCCGA 1288
179 AAGACTTTGCGCTTTTGGAGCGCGGATGCTGATTTAGTAAGTGTGTGGGATTAAGGCGCT 238
1287 AGGGCTTGGCGCTACTGAGACGGCGGATGTCAGATTAAGTGTGTGGGATTAAGGCGCC 1228
229 ACCAAGGAGAGATCAGATAGTGTGTGAGAGAGACACGACCACTGGGACTGAGACA 298
1227 ACCAAGCCTGCGATGTAGCTGTGTGAGAGAGATGATGACACATCGGAGCTGAGACA 1168
299 CGGCCCCAGACTCTTACGGGAGGAGCAGAGTGGGAAATTTTGGACAATGGGCGCAACCTGA 358
1167 CGGCCCCGAACTCTTACGGGAGGAGCAGAGTGGGAAATTTTGGACAATGGGCGCAACCTGA 1108
359 TCCAGCAATGCCCGCTGAGTGAAGAGGCTTCCGGGTTGTAAAGCTCTTTTCAGTCAAGAA 418
1107 TCCAGCCATGCGCGCTGACGAGATGAAGGCTTCCGGGTTGTAAAGCTCTTTTCAGTCAAGAA 1048
419 GAAAGGTTACGGTAAATATGTCATCTGATCGGATTCGACAGAGAGACCGGCTA 478
1047 GAAAGGCTCTTTCTTAATAAAGAGGCACTGACCGTACCGTAAAGAAATTAAGCAGCGGCTA 988
479 ACTAGCTGCGCAGACCGCGGATATACGATGAGGAGTCAAGCGTAAATCCGAATTAATCGGAC 538
987 ACTAGCTGCGCAGACCGCGGATATACGATGAGGAGTCAAGCGTAAATCCGAATTAATCGGAC 928
539 GTAAAGGCTGCGCAGCGGCTTTGTAGTCAAGTGTAAATCCCGGCTTAATCTGAGGA 598
927 GTAAAGGCTGCGCAGCGGCTTTGTAGTCAAGTGTGTAAATCCCGGCTTAATCTGAGGA 868
559 ATTGGGTTGAAATCTAAGGCTTAAGTGTGAGAGAGAGGATGGAATTCATGTGTAGC 658
867 ACTGCAATGTGACTGCAAGGCTGAGATGCGCAAGGAGGAGTGAATTCGCGGTAGC 808
659 AGTGAATGCGTAGAGATATGGAAGAACTCGATGGCGAGGACGCTCTGGGTTAAACA 718
807 AGTGAATGCGTAGAGATATGCGAGAGAACCGATGGCGAGGACGCTCTGGGTTAGCAG 748
719 CTGACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATCCAG 778
747 CTGACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGTGATCCAG 688
QY CCTTAAAGGATGCAACTAGTGTGGGCTTATAGCTGTGTAAACGAAGTACGGGT 838
DB CCTTAAAGGATGCAACTAGTGTGGGATTTATTTCTTCAATGACGAAGCTAACGCGT 628
QY GAAAGTTGACCGGCTGGGAGTACGGTCCGAGATTTAAACTCAAGAGATTGACGGGAC 898
DB GAAAGTTGACCGGCTGGGAGTACGGTCCGAGATTTAAACTCAAGAGATTGACGGGAC 568
QY CCGCAAGACGGGTGATGATGTGTTAATTCGATGCAACGCGAATAACCTTACCACCT 508
DB CCGCAAGACGGGTGATGATGTGTTAATTCGATGCAACGCGAATAACCTTACCACCT 508
QY TTGACATGTAGGAAATTTTCTTGAAGATAGTATGCT---TCGGGAACGCTTAACACAG 1015
DB TTGACATGTAGGAAATTTTCTTGAAGATAGTATGCT---TCGGGAACGCTTAACACAG 448
QY TGCTGCATGCGCTGTGTCAGCTCGTGTCTGTGAGATGTTGGGTTAAGTCCCGAAGAGAG 1075
DB TGCTGCATGCGCTGTGTCAGCTCGTGTCTGTGAGATGTTGGGTTAAGTCCCGAAGAGAG 388
QY CAACCTTTGCTATTAATTCGCATCATTTGTTGGGCACTTAAATGAGATCGCGGTGACA 1135
DB CAACCTTTGCTATTAATTCGCATCATTTGTTGGGCACTTAAATGAGATCGCGGTGACA 333
QY AACCGGAGAGAGTGGGATGACGTCATGCTCTCATGCGCTTAAAGGAGGCTTTCACA 1195
DB AACCGGAGAGAGTGGGATGACGTCATGCTCTCATGCGCTTAAAGGAGGCTTTCACA 273
QY CGTAAATACATGCGCGCTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAA 1255
DB CGTAAATACATGCGCGCTACAGAGGTTGCCAACCCGCGAGGGGAGCTAATCTCAGAAA 1255
```

```

272 CGTCATACATGCGCGGTACAAAGGAGCAGAACCCGCGAGGTGAAGCAATCCCATMAA 213
QY GCGGCTGTAGTCCCGATTCGAGTCTGCAATTCGACTCCGGAAGTCCGTAATCCCTGTA 1315
DB GCCGCTGTAGTCCCGATTCGAGTCTGCAATTCGACTCCGGAAGTCCGTAATCCCTGTA 153
QY ATCCGAGTACAGATGTCGCGGATTAATCGTCCCGGCTTTGTACACACCCCGCTTAC 1375
DB ATCCGAGTACAGATGTCGCGGATTAATCGTCCCGGCTTTGTACACACCCCGCTTAC 93
QY ACCATGGAAGTGGGTTTCCACCAAGAGAGTGTCTTAAACGTTAAGAGAGGCGCTTGCCAC 1435
DB ACCATGGAAGTGGGTTTCCACCAAGAGAGTGTCTTAAACGTTAAGAGAGGCGCTTGCCAC 33
QY CGTGAGATTCACTGCTGGGTG 1457
DB CGTGAGATTCACTGCTGGGTG 11

RESULT 60
US-10-464-356-2/c
; Sequence 2, Application US/10464356
; Publication No. US20030207320A1
; GENERAL INFORMATION:
; APPLICANT: Bramer, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; FILE REFERENCE: BCI03 US NA
; CURRENT FILING DATE: 2003-06-17
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 50/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Brachymonas
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain NBE13
US-10-464-356-2

Query Match 74.1%; Score 1079.6; DB 6; Length 1467;
Best Local Similarity 85.6%; Pred. No. 1.3e-277;
Matches 1251; Conservative 0; Mismatches 159; Indels 12; Gaps 4;

1 ATTGAACGCTGGCGGATGCTTTACATGCAATGCCAAGCGGACGACCGATGCTTGAT 60
DB 1465 ATTGAACGCTGGCGGATGCTTTACATGCAATGCCAAGCGGACGCGG--TCTTCCGG 1408
QY CTGGTGGGAGTGGCGGACGCGGTGATGTAATCATTCGGAACCTATTCAGAAAGGCGGTA 120
DB 1407 ATGCGCGGAGTGGCGGACGCGGTGATGTAATCATTCGGAACCTATTCAGAAAGGCGGTA 1348
QY 121 ACGCATGAAAGATGTGCTAATACCGCATATATCTTAAGAGGAAAGCAGGGA--TCGA 178
DB 1347 GCTCGGCGAAAGCCGATTAATACCGCATATATCTTAAGAGGAAAGCAGGGA--TCGA 1288
QY 179 AAGACCTTGGCGCTTTTGAAGCGGCGGATGTCTGATTTAGTATGTTGGGTTAAAGCGCT 238
DB 1287 AGGGCTTGGCGCTTTTGAAGCGGCGGATGTCTGATTTAGTATGTTGGGTTAAAGCGCT 1228
QY ACCAAGGAGAGATCAAGTGTGTGAGAGAGCAGACCAACCACTGGGACTGAGACA 298
DB 1227 ACCAAGGAGAGATCAAGTGTGTGAGAGAGCAGACCAACCACTGGGACTGAGACA 1168
QY CGGCCCCGAACTCTTACGGGAGGAGCAGAGTGGGAAATTTTGGACAATGGGCGCAAGCTTGA 358
DB CGGCCCCGAACTCTTACGGGAGGAGCAGAGTGGGAAATTTTGGACAATGGGCGCAAGCTTGA 1108
```

QY 359 TCACAGTAATGCCGCTGAGTGAAGAGGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAA 418  
DB 1107 TCACACCAATGCCGCTGAGTGAAGAGGCTTCGGGTTGTAAAGCTCTTTCAGTCGAGAA 1048  
QY 419 GAAAAGTTACGGTAATATATGCTGATGATGACGCTATGACAGAAAGACCGGCTTA 478  
DB 1047 GAAAAGGCTCTTCTTAATTAAGAGGACATGACGCTATCCGTAAGATTAAGACCGGCTTA 988  
QY 479 ACTAGTCGACAGCGCGGCTAATACGTAGGAGTGCAGCGTAAATCCGAAATTAATGAGC 538  
DB 987 ACTAGTCGACAGCGCGGCTAATACGTAGGAGTGCAGCGTAAATCCGAAATTAATGAGC 928  
QY 539 GTAAGAGTGCAGCGCGGCTTGTAAAGTGAATGTGAATCCCGGGCTTAACCTGAGGA 598  
DB 927 GTAAAGCGTGCAGCGCGGCTTGTGAAGACGTGTAATCCCGGGCTTAACCTGAGGA 868  
QY 599 ATTGCGTTGAAACTACAGGCTAGAGTGTGCGAGAGGAGGTGGAATTCATGTGTAC 658  
DB 867 ACTGCAATGTGATGCTCAAGGCTGAGTACGCGCAGAGGGGAGTGAATTCGCGTGTAC 808  
QY 659 AGTGAATGCGCTAGAGTATGAGAAACATGATGCGAAGCGAGCGCTCGGCTTAACA 718  
DB 807 AGTGAATGCGCTAGAGTATGAGAAACATGATGCGAAGCGAGCGCTCGGCTTAACA 748  
QY 719 CTGACGCTCATGCAACGAAAGCGTGGGAGCAACAGAAATTAATCCTGTAGTCACG 778  
DB 747 CTGACGCTCATGCAACGAAAGCGTGGGAGCAACAGAAATTAATCCTGTAGTCACG 688  
QY 779 CCCTTAACGATGCACTAGTGTGGGCTTATTAAGGCTGTAACGAACTAACGCGT 838  
DB 687 CCCTTAACGATGCACTAGTGTGGGCTTATTAAGGCTGTAACGAACTAACGCGT 628  
QY 839 GAAGTTGACCGCTGGGAGTACGCTGCGAAGTAAATCTCAAGAAATGACGGGAGC 898  
DB 627 GAAGTTGACCGCTGGGAGTACGCTGCGAAGTAAATCTCAAGAAATGACGGGAGC 568  
QY 899 CCGCAACAGCGGTGATATGATGATTAATTCGATCAACGCGAAATCCTTAACCTACC 958  
DB 567 CCGCAACAGCGGTGATATGATGATTAATTCGATCAACGCGAAATCCTTAACCTACC 508  
QY 959 TTGACATGATGCGAATTTCTAGAGATGATTAAGTCT--TCGGGAACGCTTAACAGC 1015  
DB 507 TTGACATGATGCGAATTTCTAGAGATGATTAAGTCT--TCGGGAACGCTTAACAGC 448  
QY 1016 TGTGTGATGCTGTGCTGACGCTGCTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGC 1075  
DB 447 TGTGTGATGCTGTGCTGACGCTGCTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGC 388  
QY 1076 CAACCTGTGCTTAATTTGCGATCATTTGGTGGGCACTTAAATGAGATGCGCGTGA 1135  
DB 387 CAACCTGTGCTTAATTTGCGATCATTTGGTGGGCACTTAAATGAGATGCGCGTGA 333  
QY 1136 AACCGAGGAAGGTGGGATGAGCTCAAGTCTCATGAGCCCTTAATGGGATGGGCTTACA 1195  
DB 332 AACCGAGGAAGGTGGGATGAGCTCAAGTCTCATGAGCCCTTAATGGGATGGGCTTACA 273  
QY 1196 CGTAATCAATGCGCGCTACAGAGGTTGCCAACCGCGAGGGGAGCTTAATCTGAGAA 1255  
DB 272 CGTAATCAATGCGCGCTACAGAGGTTGCCAACCGCGAGGGGAGCTTAATCTGAGAA 213  
QY 1256 GCGCGTGTAGTCCGAGTCCGAGTCTGCAACTGCACTCCGTAAGTGGAGTCCGCTGTA 1315  
DB 212 GCGCGTGTAGTCCGAGTCCGAGTCTGCAACTGCACTCCGTAAGTGGAGTCCGCTGTA 153  
QY 1316 ATGCGGATGACATGTCGCGGTGAATACGTTCCCGGCTTGTGACACACCGCCGTAC 1375  
DB 152 ATGCGGATGACATGTCGCGGTGAATACGTTCCCGGCTTGTGACACACCGCCGTAC 93  
QY 1376 ACCATGGAGTGGTTTCAACAGAGAGAGTCTAAACGTTAGAGAGGGGCTTGGCAC 1435  
DB 92 ACCATGGAGTGGTTTCAACAGAGAGAGTCTAAACGTTAGAGAGGGGCTTGGCAC 33

QY 1436 GGTGATTCATGACTGGGTTG 1457  
DB 32 GGTGGGTTCTGACTGGGTTG 11  
RESULT 61  
US-10-464-709-2/c  
; Sequence 2, Application US/10464709  
; Publication No. US20030207321A1  
; GENERAL INFORMATION:  
; APPLICANT: Brumicci, Michael  
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in  
; TITLE OF INVENTION: Industrial Wastewater Bioreactors  
; FILE REFERENCE: BC1033 US NA  
; CURRENT APPLICATION NUMBER: US/10/464,709  
; PRIOR FILING DATE: 2003-06-17  
; PRIOR APPLICATION NUMBER: US/09/735,567  
; PRIOR FILING DATE: 2000-12-13  
; PRIOR APPLICATION NUMBER: 60/171,140  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 1467  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Brachymonas  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain NBE13  
US-10-464-709-2  
Query Match 74.1%; Score 1079.6; DB 6; Length 1467;  
Best Local Similarity 85.6%; Pred. No. 1.3e-277;  
Matches 1251; Conservative 0; Mismatches 199; Indels 12; Gaps 4;  
QY 1 ATGGAACGTTGCGCGGATGCTTATCAATGCAAGTGAACGCGAGCAGGATCTTGCAAT 60  
DB 1465 ATGGAACGTTGCGCGGATGCTTATCAATGCAAGTGAACGCGAGCAGGATCTTGCAAT 1408  
QY 61 CTGTTGGCAGTGGCGGAGCGGCTGATATGATCGAACTATCCAGAAAGGGGGTGA 120  
DB 1407 ATGCGCGGAGTGGCGGAGCGGCTGATATGATCGAACTATCCAGAAAGGGGGTGA 1348  
QY 121 ACGCATGAAAGATGTGCTAATACGCAATATCTTAAGAGAGAAAGCGGGA--TCGA 178  
DB 1347 GCTCGGCAAAAGCGGATTAATACGCAATATCTTAAGAGAGAAAGCGGGA--TCGA 1288  
QY 179 AAGACTTGGCGCTTTTGGAGCGCGCGATGCTGATTAAGTGTGGGTTAAAGGCTT 238  
DB 1287 AAGACTTGGCGCTTTTGGAGCGCGCGATGCTGATTAAGTGTGGGTTAAAGGCTT 1228  
QY 239 ACCAAGGCGCATGATGATGTTGCTGAGAGGACAGCAGCAGCACTGGACTGAGACA 298  
DB 1227 ACCAAGGCGCATGATGATGTTGCTGAGAGGATATGAGCAGCAGCACTGGACTGAGACA 1168  
QY 299 CGGCGCAGACTCTTAACGGAAGCGAGCAGATGGGGAATTTGACAATGGCGCAAGCTGA 358  
DB 1167 CGGCGCAGACTCTTAACGGAAGCGAGCAGATGGGGAATTTGACAATGGCGCAAGCTGA 1108  
QY 359 TCAGCAATGCGCGCTGATGAGAAAGGCTTGGGTTGTAAGCTCTTCACTCGAGAA 418  
DB 1107 TCAGCAATGCGCGCTGATGAGAAAGGCTTGGGTTGTAAGCTCTTCACTCGAGAA 1048  
QY 419 GAAAAGTTACGGTAATATATGCTGATGATGACGCTATGACAGAAAGACCGGCTTA 478  
DB 1047 GAAAAGGCTCTTCTTAATTAAGAGGACATGACGCTATCCGTAAGATTAAGACCGGCTTA 988  
QY 479 ACTAGTCGACAGCGCGGCTAATACGTAGGAGTGCAGCGTAAATCCGAAATTAATGAGC 538  
DB 987 ACTAGTCGACAGCGCGGCTAATACGTAGGAGTGCAGCGTAAATCCGAAATTAATGAGC 928  
QY 539 GTAAGAGTGCAGCGCGGCTTGTAAAGTGAATGTGAATCCCGGGCTTAACCTGAGGA 598

||||| 868  
Db 927 GTAAGCGGCCAGCGCTCTTGCAGACAGTGGTGAATCCCGGGCTCAACTGGGA  
Qy 599 ATTGCGTTGAACATCAAGAGCTAGGTGCGAGAGGAGTGGATTCATGTGTAC  
Db 867 ACTGCATATGTACTGCAAGGCTGGGTACGGCAGAGGGGATGGAATTCGCCGTGAC  
Qy 659 AGTGAATGCGTAGATATGGAAGACATCGATGGCGAAGGACCTCTGGTTACA  
Db 807 AGTGAATGCGTAGATATGCGAGAGACCCATGGCGAAGGACGTCCTTGGGCTGTA  
Qy 719 CTGAGCTCATGCAAGAAAGCGTGGGAGCAACAGATTAGATACCTGTATGCAAG  
Db 747 CTGAGCTCATGCAAGAAAGCGTGGGAGCAACAGATTAGATACCTGTATGCAAG  
Qy 779 CCTTAACGATGCACTGATGTTGGCTTATTAAGCTTGTATGCAAGTAAAGCGT  
Db 687 CCTTAACGATGCACTGATGTTGGGATTTATTTCTTCACTAAGCAAGCTAAAGCGT  
Qy 839 GAACTTGAACCGCTGGGGAGTACGGTCCGCAAGATTAAACCTAAGAAATTGACGGGAC  
Db 627 GAACTTGAACCGCTGGGGAGTACGGTCCGCAAGATTAAACCTAAGAAATTGACGGGAC  
Qy 899 CCGCAAGAGCGGTGATTTATGTGATTAATTCGATGCAACGCAAAAACCTTACCTAC  
Db 567 CCGCAAGAGCGGTGATTTATGTGATTAATTCGATGCAACGCAAAAACCTTACCTAC  
Qy 959 TTGACATGATGCAATTTTCTAGAGATGATTAAGTCT--TCGGAAACCTTACACAG  
Db 507 TTGACATGATGCAATTTTCTAGAGATGATTAAGTCT--TCGGAAACCTTACACAG  
Qy 1016 TGCTCATGAGCTGTGTCAGCTCGTGTGATGATGTTAGTCCCGCAAGAGG  
Db 447 TGCTCATGAGCTGTGTCAGCTCGTGTGATGATGTTAGTCCCGCAAGAGG  
Qy 1076 CAACCTTGTCATTAATTCGATTTGGTGGGCACTTATGAGACTGCGGCTGACA  
Db 387 CAACCTTGTCATTAATTCGATTTGGTGGGCACTTATGAGACTGCGGCTGACA  
Qy 1136 AACCGAGAGAGTGGGATGACGTCAGTCTTACGCTTATGAGAGGCTTACCA  
Db 332 AACCGAGAGAGTGGGATGACGTCAGTCTTACGCTTATGAGAGGCTTACCA  
Qy 1196 CGTATTAAGTGGCGGTACAGAGGTTGCCAACCGGAGGGAGCTTATCTAGAA  
Db 272 CGTATTAAGTGGCGGTACAGAGGTTGCCAACCGGAGGGAGCTTATCTAGAA  
Qy 1256 GCGCTCGTAGTCCCGATCGAGTCTGCAACTGCACTCGGTGAGTCCGGAATCGCTAGTA  
Db 212 GCGCTCGTAGTCCCGATCGAGTCTGCAACTGCACTCGGTGAGTCCGGAATCGCTAGTA  
Qy 1316 ATTCGCGATCAGCATGTCGGGTGAATACGTTCCGGGTCTTTTACACACCGCCCTGAC  
Db 152 ATTCGCGATCAGCATGTCGGGTGAATACGTTCCGGGTCTTTTACACACCGCCCTGAC  
Qy 1376 ACCATGGAGTGGGTTTACACGAGCAGGTGATACCGTAAAGAGGGGCTTGGCAC  
Db 92 ACCATGGAGTGGGTTTACACGAGCAGGTGATACCGTAAAGAGGGGCTTGGCAC  
Qy 1436 GGTGATTCATGACTGGGGTG 1457  
Db 32 GGTGGGTTCTGTGACTGGGGTG 11

RESULT 62  
US-10-168-337A-7  
; Sequence 7, Application US/10168337A  
; Publication No. US20030170654A1  
; GENERAL INFORMATION:  
; APPLICANT: Crocetti, Gregory R.; Tyson, Gene W.; Hugenholtz, Philip; and Blackall,  
; TITLE OF INVENTION: Probes and Primers for the Detection of Polyporphate  
; TITLE OF INVENTION: Accumulating Organisms in Wastewater

FILE REFERENCE: 00267PC/KF  
; CURRENT APPLICATION NUMBER: US/10/168,337A  
; NUMBER OF SEQ IDS: 14  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1320  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:  
US-10-168-337A-7

Query Match 74.1%; Score 1079.2; DB 6; Length 1320;  
Best Local Similarity 90.0%; Pred. No. 1.7e-277;  
Matches 1190; Conservative 1; Mismatches 124; Indels 7; Gaps 3;

Qy 1 ATTGAACGCTGGCGCATGCTTTACATGCAAGTCCGAACCGGACGACGAGTGGTGCAT 60  
Db 1 ATTGAACGCTGGCGCATGCTTTACATGCAAGTCCGAACCGGACGAGTGGGCAAC-- 58  
Qy 61 CTGGTGGAGTGGCGGACGGGTGATGATGATCGGAACGTAATCCAGAAAGGGGGTGA 120  
Db 59 CTGGTGGAGTGGCGGACGGGTGATGATGATCGGAACGTAATCCAGAAAGGGGGTGA 118  
Qy 121 ACGATGGAAGAGTGTCTATACCGCATATCTTAAGAGGAAACGAGGGGATTCGAAA 180  
Db 119 ACGAGGAAAGCTACGCTAATACCGCATATCTTAAGAGGAAACGAGGGGATTCGAAA 178  
Qy 181 GACCTTGCTTTTGAAGCGCGCGATGCTGATTAAGTGGTGGGTTAAAGCTTAC 240  
Db 179 GACCTTGCTTTTGAAGCGCGCGATGCTGATTAAGTGGTGGGTTAAAGCTTAC 238  
Qy 241 CAAGCGACATGATGATGTTGTTGAGAGGACGACCACTGAGACTGAGACAG 300  
Db 239 CAAGCGACATGATGATGTTGTTGAGAGGATGATCCGCACTGAGACTGAGACAG 298  
Qy 301 GCCCGACTCTTACGAGAGGACGATGGGAAATTTTGAACAAAGGCGCAAGCTGATC 360  
Db 299 GCCCGACTCTTACGAGAGGACGATGGGAAATTTTGAACAAAGGCGCAAGCTGATC 358  
Qy 361 CAGCAATGCGCGTGAAGGAAAGGCTTGGGTTTAAAGCTTTCAGTCCAGAGA 420  
Db 359 CAGCAATGCGCGTGAAGGAAAGGCTTGGGTTTAAAGCTTTCAGTCCAGAGA 418  
Qy 421 AAAGTTACGTTAATATATGTCATGATCGGTATCGACAGAAAGCACCGGCTAAC 480  
Db 419 AAAGTTACGTTAATATATGTCATGATCGGTATCGGTATCGGAAATCGGCTAAC 478  
Qy 481 TACGTGCAAGAGCGCGGTAAATACGTAGGTGCAACCGTTAAATCGAAATTCAGGCGT 540  
Db 479 TACGTGCAAGAGCGCGGTAAATACGTAGGTGCAACCGTTAAATTCAGGCGT 538  
Qy 541 AAAGGTGCGGACGCGCTTTTGAATGATGATGTAATCCCGGGCTTAACTGGGAAT 600  
Db 539 AAAGGTGCGGACGCGCTTTTGAATGATGATGTAATCCCGGGCTTAACTGGGAAT 598  
Qy 601 TGCGTTGAACATCAAGAGCTAGAGTGGCAGAGGAGGTGGAATTCATGTGTACAG 660  
Db 599 TGCAATTTGAGACTCAAGACTGAGTGGCAGAGGAGGTGGAATTCATGTGTACAG 658  
Qy 661 TGAATTCGTTAAGATATGGAAGAACTGATGGCGAAGGACGCTCTGGGTTAAACT 720  
Db 659 TGAATTCGTTAAGATATGGAAGAACTGATGGCGAAGGACGCTCTGGGTTAAACT 718  
Qy 721 GACGCTATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGTATGTCACGCGC 780  
Db 719 GACGCTATGACGAAAGCGTGGGAGCAACAGATTAGATACCTGTATGTCACGCGC 778  
Qy 781 CTAACGATGCACTGATGTTGG--GCTTATTAAGCTTGGTGAACGAAGCTAACCGGT 838  
Db 779 CTAACGATGCACTGATGTTGGGAGGTTAACTTTAGTGGCGTGAACCGGT 836

839 GAAGTTGACCGGCTGGGAGTACGGTCCGAGATTAAACTCAAGAAATTGACGGGAC 898  
 839 GAAGTTGACCGGCTGGGAGTACGGTCCGAGATTAAACTCAAGAAATTGACGGGAC 898  
 899 CCGCAAGACGGGTGATTTATGTGATTTAATTCATGCAACGGGAAAACTTACCTACCC 958  
 899 CCGCAAGACGGGTGATTTATGTGATTTAATTCATGCAACGGGAAAACTTACCTACCC 958  
 959 TTGAATGATGAGGAATTTTCTAGAGATAGATTAGTCT--TCGGGAACGCTACACAGG 1015  
 959 TTGAATGATGAGGAATTTTCTAGAGATAGATTAGTCT--TCGGGAACGCTACACAGG 1018  
 1016 TGCTGATGCTGTGTGCTGAGCTGTGTGCTGAGATGTTGGGTTAACTCCGCAACGAGG 1075  
 1019 TGCTGATGCTGTGTGCTGAGCTGTGTGCTGAGATGTTGGGTTAACTCCGCAACGAGG 1078  
 1076 CAACCTTGCTATTAATTCATCATTTGTTGGGCACTTTAATGAGACTGCCGCTGACA 1135  
 1079 CAACCTTGCTATTAATTCATCATTTGTTGGGCACTTTAATGAGACTGCCGCTGACA 1138  
 1136 AACCGGAGAGGTGGGATGAGCTGAGTCCCTGATGAGGCTTATGAGGCTTACACA 1195  
 1139 AACCGGAGAGGTGGGATGAGCTGAGTCCCTGATGAGGCTTATGAGGCTTACACA 1198  
 1196 CGTAATACATGCGCGGTACAGAGGTTGCGAACCCGCGAGGGGAGCTAATCTCAGAAA 1255  
 1199 CGTATACATGCGCGGTACAGAGGTTGCGAACCCGCGAGGGGAGCTAATCTCAGAAA 1258  
 1256 GCGCGTGTGATTCGGATCGAGTCTGCAACTGACTCCGTGAGTGGAAATGCTAGTAA 1315  
 1259 GCGCGTGTGATTCGGATCGAGTCTGCAACTGACTCCGTGAGTGGAAATGCTAGTAA 1318  
 1316 AT 1317  
 1319 AT 1320

RESULT 63

US-09-985-846-1  
 / Sequence 1, Application US/09985846  
 / Publication No. US20030165471A1

GENERAL INFORMATION:

APPLICANT: Casida, Jr., Lester B.  
 Falkinham, III, Joseph Oliver

TITLE OF INVENTION: NON-OBLIGATE PREDATORY BACTERIUM  
 BURKHOLDERIA CASIDAE AND USES THEREOF

NUMBER OF SEQUENCES: 1  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP  
 STREET: 1667 K Street, N.W.  
 CITY: Washington  
 STATE: DC

COUNTRY: USA  
 ZIP: 20006

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/985 846  
 FILING DATE: 06-NO. US20030165471A1-2001

CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/063, 898  
 FILING DATE: 22-APR-1998

APPLICATION NUMBER: 60/044, 532  
 FILING DATE: 23-APR-1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Mirock, Leslie  
 REGISTRATION NUMBER: 18, 872

REFERENCE/DOCKET NUMBER: 8743-006-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 7509090  
 TELEFAX: (212) 8699741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1495 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: Genomic DNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-985-846-1

Query Match 72.8%; Score 1060.8; DB 3; Length 1495;  
 Best Local Similarity 86.6%; Pred. No. 1.4e-272;  
 Matches 1264; Conservative 0; Mismatches 182; Indels 14; Gaps 8;

1 ATTGAACGCTGGCGGCAATGCTTTACA-CATGCAATGCAACGGGACGAGGCTTGA 59  
 3 ATATTACGCTGCTGCTGCTTACATGCTTACATGCAATGCAACGGGACGAGGCTTGA 62  
 60 TCTGGTGGGAGTGGCGGAGCGGCTGATGATGATGATGATGATGATGATGATGATGAT 118  
 63 CTTGGTGGGAGTGGCGGAGCGGCTGATGATGATGATGATGATGATGATGATGATGAT 122  
 119 TAACCATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 178  
 123 TAGCCCGGAGAAAGCGGAGTAAATCCGATACGATGATGATGATGATGATGATGATGAT 182  
 179 AAGACCTGCGCTTTTGGAGCGGCGGATGATGATGATGATGATGATGATGATGATGAT 238  
 183 CGGCGCTCGCGCTATGAGGCTGCGGATGATGATGATGATGATGATGATGATGATGAT 242  
 239 ACCAAGGCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298  
 243 ACCAAGGCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301  
 299 CGGCGCGAGCTCTTACGAGGAGCGAGTGGGAAATTTTGAACAATGGCGGCAAGCTGA 358  
 302 CGGCGCGAGCTCTTACGAGGAGCGAGTGGGAAATTTTGAACAATGGCGGCAAGCTGA 361  
 359 TCCAGCAATGCGCGGCTGATGAAAGGCTTGGGTTGTAAGCTCTTCACTCGAGAA 418  
 362 TCCAGCAATGCGCGGCTGATGAAAGGCTTGGGTTGTAAGCTCTTCACTCGAGAA 421  
 419 GAAAGCTTACGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 478  
 422 GAAATCTTGTGTTCTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481  
 479 ACTACGTGCGAGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538  
 482 ACTACGTGCGAGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541  
 539 GTAAAGGCTGCGAGCGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 598  
 542 GTAAAGGCTGCGAGCGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 601  
 599 ATTGCGTTGAAATCTAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 658  
 602 ACTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661  
 659 AGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
 662 AGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721  
 719 CTGACGCTATGACGAAAGCGTGGGAGCAACAGATTTAGTACCTCTGTGTGTCCAGC 778  
 722 CTGACGCTATGACGAAAGCGTGGGAGCAACAGATTTAGTACCTCTGTGTGTCCAGC 781  
 779 CCTTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838

Db 782 CCTAAGCATGTCAACTAGTTGTTGGGATTCATTCTCTAGTAACGTAAACGGCT 841  
Qy 839 GAAAGTTGACCCGCTGGGAGTACGGTCCGAAGATTAAATCAAGGAATTGACGGGAC 898  
Db 842 GAAAGTTGACCCGCTGGGAGTACGGTCCGAAGATTAAATCAAGGAATTGACGGGAC 901  
Qy 899 CCGCAAGAGCGGTGATTTATGATTAATTTGATGCAACGGGAAAACCTTACCTACCC 958  
Db 902 CCGCAAGAGCGGTGATTTATGATTAATTTGATGCAACGGGAAAACCTTACCTACCC 961  
Qy 959 TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTCT--TCGGGAACGCTAACACAG 1015  
Db 962 TTGACATGTAGCGAATTTTCTAGAGATAGATTAGTCTGAAAAGAAACCGGACAGG 1021  
Qy 1016 TGCTGATGCTGTGTCTGCTGCTGTCTGTGAGATTTGGGTTAAAGTCCCGCAAGAGG 1075  
Db 1022 TGCTGATGCTGTGTCTGCTGCTGTCTGTGAGATTTGGGTTAAAGTCCCGCAAGAGG 1081  
Qy 1076 CAACCTTGCTATTAATTTGATCAATTTGTTGGGCACTTAATGAGACTGCGGTGACA 1135  
Db 1082 CAACCTTGCTATTAATTTGATCAATTTGTTGGGCACTTAATGAGACTGCGGTGACA 1136  
Qy 1136 AACCGAGAGAGTGGGAGTGAAGTCAAGTCTCAATGAGCTTATGGGTAGGGCTTACA 1195  
Db 1137 AACCGAGAGAGTGGGAGTGAAGTCAAGTCTCAATGAGCTTATGGGTAGGGCTTACA 1195  
Qy 1196 CGTATATCAATGGCGGTACAAAGGTTGCCAACCCCGAGAGGGAGCTTATCTCAAGAA 1255  
Db 1196 CGTATATCAATGGCGGTACAAAGGTTGCCAACCCCGAGAGGGAGCTTATCTCAAGAA 1254  
Qy 1256 GCGCGTGTAGTCCGGATCGAGTCTGCAACTGCACTCCGTAAGTCCGATCGTAGTA 1315  
Db 1255 ACCGATCGTAGTCCGGATCGAGTCTGCAACTGCACTCCGTAAGTCCGATCGTAGTA 1314  
Qy 1316 ATCGCGATCAGCATGTGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCGTAC 1375  
Db 1315 ATCGCGATCAGCATGTGCGGTGAATACGTTCCCGGCTTTGTACACACCGCCGTAC 1374  
Qy 1376 ACCAAGGAGAGTGGTTTACCAAGAGCAGTAGTCTTAACCGTAAGG-AGGGGCTTGGCA 1434  
Db 1375 ACCAAGGAGAGTGGTTTACCAAGAGTGGTTTACCAAGAGTGGTTTACCAAGAGAGGCTCCCA 1434  
Qy 1435 CGGTGAGATTGACTGGG 1454  
Db 1435 CGGTGAGATTGACTGGG 1454

RESULT 64  
US-10-776-767-1  
Sequence 1, Application US/10776767  
Publication No. US20040127687A1  
GENERAL INFORMATION:  
APPLICANT: Casida, Jr., Lester E.  
Falkinham, III, Joseph Oliver  
Cain, Cody C.  
TITLE OF INVENTION: NON-OBLIGATE PREDATORY BACTERIUM  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1667 K Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/776,767  
FILING DATE: 10-Feb-2004

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/985,846  
FILING DATE: 06-Nov-2001  
APPLICATION NUMBER: 09/063,898  
FILING DATE: 22-APR-1998  
APPLICATION NUMBER: 60/044,532  
FILING DATE: 23-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mirock, Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 8743-006-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 7909090  
TELEFAX: (212) 8699741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: genomic DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-776-767-1

Query Match 72.8%; Score 1060.8; DB 7; Length 1495;  
Best Local Similarity 86.6%; Pred. No. 1.4e-272;  
Matches 1264; Conservative 0; Mismatches 162; Indels 14; Gaps 8;

Qy 1 ATTTAAGCTGTGGCGGATGCTTTTACA-CATGCAAGTGAACCGGACGAGCTTGGCA 59  
Db 3 ATTTAAGCTGTGGCGGATGCTTTTACAAGATGCAAGTGAACCGGACGAGCTTGGCA 62  
Qy 60 TCTGTGCGAGATGCGGACCGGTGAGTATGATCGGAAC-GTATCCAGAAAGGGGG 118  
Db 63 CTTGTGCGAGATGCGGACCGGTGAGTATGATCGGAAC-GTATCGGAAGGGGG 122  
Qy 119 TAAAGCATCGAAGATGCTTAATACGCAATATCTTAAGAGAGAAAGCGGGGATGCA 178  
Db 123 TAGCCCGGAGAAACCGGATTAATACGCAATATCTTAAGAGAGAAAGCGGGGATGCA 182  
Qy 179 AAGACCTTGGCGCTTTTGGAGCGGCGGATGCTGATTAAGTGTGTGGGTAAGGCTT 238  
Db 183 CGGCGCTGCGCTTATAGGTTGGCGGATGCTGATTAAGTGTGTGGGTAAGGCTT 242  
Qy 239 ACCAAGCGAGATGATGTTGCTGAGAGAGAGACAGCCACACTGGGACTGAGACA 298  
Db 243 ACCAAGCGAGATGATGTTGCTGAGAGAGAGAGACAGCCACACTGGGACTGAGACA 301  
Qy 299 CCGGCCAGACTCTTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358  
Db 302 CCGGCCAGACTCTTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361  
Qy 359 TCCAGCAATGCGCGGTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418  
Db 362 TCCAGCAATGCGCGGTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421  
Qy 419 GAAAGGTTACGTAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 478  
Db 422 GAAATCTTGTTGTTAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 481  
Qy 479 ACTAAGTGCAG 538  
Db 482 ACTAAGTGCAG 541  
Qy 539 GTAAGGTTGCGAG 598  
Db 542 GTAAGGTTGCGAG 601  
Qy 599 ATTGCGTTTGAAGTCAAG 658  
Db 602 ACTGCAATGTAAGTCAAG 661



QY 895 GGAACCCGCAAGAGGGTGGATTATGATGATTAATTCGATGCAAGCGAAGAACTTACT 954  
 DB 924 GGGCCGCAAGAGGGTGGATTATGATGATTAATTCGATGCAAGCGAAGAACTTACT 983  
 QY 955 ACCCTTGACATGACGGAATTTTCTAGATAGATTAGT-CTTGGGAAAGCTTAACA 1013  
 DB 984 GGCCTTGACATGACGGAATTTTCTAGATAGATTAGT-CTTGGGAAAGCTTAACA 1043  
 QY 1014 GGTGCTGATGAGCTGTCTGACGTGTGTCTGATGATGTTGGTTAACTCCCGCAAG 1073  
 DB 1044 GGTGCTGATGAGCTGTCTGACGTGTGTCTGATGATGTTGGTTAACTCCCGCAAG 1103  
 QY 1074 GCGAACCTTGATCATTAATGTCATCA--TTTGGTGGGCACTTAATGAGACTCGGAT 1131  
 DB 1104 GCGAACCTTGATCATTAATGTCATCA--TTTGGTGGGCACTTAATGAGACTCGGAT 1163  
 QY 1132 GACAAACCGAGAGAGGTGGGGATGACGTCAAGTCTCAATGACCTTATGGGTAGGGCTT 1191  
 DB 1164 GACAAACCGAGAGAGGTGGGGATGACGTCAAGTCTCAATGACCTTATGGGTAGGGCTT 1223  
 QY 1192 CACAGTATATCAATGAGCGGCTTACAGAGGTTGCCAACCCGAGAGGGAGCTATCTCA 1251  
 DB 1224 CACAGTATATCAATGAGCGGCTTACAGAGGTTGCCAACCCGAGAGGGAGCTATCTCA 1283  
 QY 1252 GAAAGCGGTGTAGTCTCGGATCGAGATCTGCACTCGTGAAGTGGAAATCGCT 1311  
 DB 1284 GAAAGCGGTGTAGTCTCGGATCGAGATCTGCACTCGTGAAGTGGAAATCGCT 1343  
 QY 1312 AGTAATCGCGATACGCA-TGTGCGGATGAAATACGTTCCCGGCTTTGACACACCGCC 1370  
 DB 1344 AGTAATCGCGATACGCA-TGTGCGGATGAAATACGTTCCCGGCTTTGACACACCGCC 1403  
 QY 1371 GTCAACCATGAGGATGGTTTACACAGAGCATGATCTTAACCTGAAGAGGCGCTT 1430  
 DB 1404 GTCAACCATGAGGATGGTTTACACAGAGCATGATCTTAACCTGAAGAGGCGCTT 1463  
 QY 1431 GCCACGGTGAATCATGACTGGGGT 1457  
 DB 1464 GCCACGGTGAATCATGACTGGGGT 1490

RESULT 66  
 US-10-260-647-1  
 ; Sequence 1, Application US/10260647  
 ; Publication No. US20030104605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WALKER, Harrell L.  
 ; APPLICANT: HIGGINBOTHAM, Lawrence R.  
 ; TITLE OF INVENTION: CONTROL OF CYANOBACTERIA WITH A BACTERIUM  
 ; FILE REFERENCE: 32158.8  
 ; CURRENT APPLICATION NUMBER: US/10/260,647  
 ; CURRENT FILING DATE: 2002-09-27  
 ; PRIOR APPLICATION NUMBER: US 09/967,376  
 ; PRIOR FILING DATE: 2001-09-28  
 ; NUMBER OF SEQ ID NOS: 1  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 1540  
 ; TYPE: DNA  
 ; ORGANISM: Unknown Organism  
 ; FEATURE:  
 ; OTHER INFORMATION: Bacterium SG-3, gram negative, rod-shaped; exhibits flagellate  
 ; OTHER INFORMATION: motility; pathogenic to cyanobacteria and algae;  
 ; OTHER INFORMATION: yellow colonies on BG-11 medium suppl. with trytic  
 ; OTHER INFORMATION: soy  
 US-10-260-647-1

Query Match 72.3%; Score 1053; DB 5; Length 1540;  
 Best Local Similarity 84.7%; Pred. No. 1.7e-270;  
 Matches 1242; Conservative 0; Mismatches 215; Indels 10; Gaps 5;  
 QY 1 ATTGAACGCTGCGCATGCTTACATGCAAGTCAAGCGGACACGAGATGCTTGCAT 60  
 DB 1 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 24 AGTGAACGCTGCGGAGGCTTAACATGCAAGTCAAGCGGACACAGCATGACAT 83  
 QY 61 CT----GTTGCGGATGCGGAGCGGATGAGTAATCGGATTCGAGAGAGGG 116  
 DB 84 ACTGCGGTGCGAGTGGCGGAGCGGATGAGTAATCGGATTCGAGAGAGGG 143  
 QY 117 GGTAAAGCATGAAAGATGCTAATACCGCATTAATCTTAAGAGAGAGAGGATC 176  
 DB 144 GATTAACATAGGAAACTTATCTAATACCGCATTAATCGGATTCGAGAGAGGATC 203  
 QY 177 GAAAGACTTTCGCTTTTGAAGCGGCGGATGCTGAATAGTACTAGTGGTGGGTAAGGC 236  
 DB 204 GAAAGACTTTCGCGGATGATGAGCGGATTCGATTAATGCTAGTGGCGGATTAAGGC 263  
 QY 237 CTACCAAGGCGAGCATGATGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 296  
 DB 264 CCACCAAGGCGAGCATGATGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 323  
 QY 297 CACGCGCAGACTCTTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 356  
 DB 324 CACGCTCAGACTCTTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 383  
 QY 357 GATCAAGCATGCGCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416  
 DB 384 GATCAAGCATGCGCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 443  
 QY 417 AAGAAAGGTTACGTTAAATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476  
 DB 444 AAGAAAGGTTACGTTAAATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 503  
 QY 477 TAACTAGTGCAG 536  
 DB 504 TAACTAGTGCAG 563  
 QY 537 GCGTAAAGGAGTGCAG 596  
 DB 564 GCGTAAAGGAGTGCAG 623  
 QY 597 GAAATGCTTTGAAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 656  
 DB 624 GAAATGCTTTGAAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 683  
 QY 657 GCAATGAAATGCGTGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 716  
 DB 684 GCAATGAAATGCGTGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 743  
 QY 717 CACTGACGCTGATGAG 776  
 DB 744 CACTGACGCTGATGAG 803  
 QY 777 GCGCTTAAAGATGATCAATGATGTTGGGCTTATTAAG--CTTGGTAAAGAGCTAAC 834  
 DB 804 GCGCTTAAAGATGATCAATGATGTTGGGCTTATTAAG--CTTGGTAAAGAGCTAAC 863  
 QY 835 GCGTGAAGTTGACCGGCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 894  
 DB 864 GCGTGAAGTTGACCGGCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923  
 QY 895 GGAACCCGCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
 DB 924 GGGCCGCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 983  
 QY 955 ACCCTTGACATGACGGAATTTTCTAGATAGATTAGT-CTTGGGAAAGCTTAACA 1013  
 DB 984 GGCCTTGACATGACGGAATTTTCTAGATAGATTAGT-CTTGGGAAAGCTTAACA 1043  
 QY 1014 GGTGCTGATGAGCTGTCTGACGTGTGTCTGATGATGTTGGTTAACTCCCGCAAG 1073  
 DB 1044 GGTGCTGATGAGCTGTCTGACGTGTGTCTGATGATGTTGGTTAACTCCCGCAAG 1103  
 QY 1074 GCGAACCTTGATCATTAATGTCATCA--TTTGGTGGGCACTTAATGAGACTCGGAT 1131  
 DB 1104 GCGAACCTTGATCATTAATGTCATCA--TTTGGTGGGCACTTAATGAGACTCGGAT 1163



QY 1132 GACAAACCGGAGGAGTGGGGATGACGTCAAGTCTCTATGCGCCCTTATGGTAGGGCTT 1191  
DB 1164 GACAAACCGGAGGAGTGGGGATGACGTCAAGTCTCTATGCGCCCTTATGGTAGGGCTT 1223  
QY 1192 CACAGCTAATACATAGGCGCGCTTACAGAGGCTTCCACCCCGGAGGGAGGATATCTCA 1251  
DB 1224 CACAGCTAATACATAGGCGCGCTTACAGAGGCTTCCACCCCGGAGGGAGGATATCTCA 1283  
QY 1252 GAAAGCGGCTGTAGTCCGATCGAGTCTGCACTCGACTCCGTAAGTCCGAATCGCT 1311  
DB 1284 GAAAGCGGCTGTAGTCCGATCGAGTCTGCACTCGACTCCGTAAGTCCGAATCGCT 1343  
QY 1312 AGTAAATCGGATCAACA-TGTGCGGCTGAATACGTTCCGGGCTTGTACACACCGCC 1370  
DB 1344 AGTAAATCGGATCAACA-TGTGCGGCTGAATACGTTCCGGGCTTGTACACACCGCC 1403  
QY 1371 GTCAACCCATGGGAGTGGGTTTCAACAGAAAGCAAGTACTTAACTGTAAGAGGGCGCTT 1430  
DB 1404 GTCAACCCATGGGAGTGGGTTTCAACAGAAAGCAAGTACTTAACTGTAAGAGGGCGCTT 1463  
QY 1431 GCCAGGTGAGATTCATGACTGGGGTG 1457  
DB 1464 GCCAGGTGAGATTCATGACTGGGGTG 1490

## RESULT 67

US-10-297-465A-1  
Sequence 1, Application US/10297465A  
General Information:  
Applicant: Simpson, Andrew  
Applicant: Reinach, Fernando  
Applicant: Setubal, Joao  
Applicant: Mediane, Paulo  
Title of Invention: Isolated Genome of *Xylella fastidiosa* and Uses Thereof  
File Reference: FAPESP 202 US (10213376)  
Current Application Number: US/10297465A  
Current Filing Date: 2001-06-07  
Prior Application Number: PCT/IB01/01618  
Prior Filing Date: 2001-06-07  
Prior Application Number: 60/209,906  
Prior Filing Date: 2001-06-17  
Number of SEQ ID NOS: 1  
Software: PatentIn version 3.2  
SEQ ID NO: 1  
Length: 2731748  
Type: DNA  
Organism: *Xylella fastidiosa*  
US-10-297-465A-1

Query Match 71.5%; Score 1041.8; DB 7; Length 2731748;  
Best Local Similarity 84.2%; Pred. No. 5.6e-267;  
Matches 1235; Conservative 0; Mismatches 222; Indels 10; Gaps 5;  
QY 1 ATTTGAACGCTGGGCGGATGCTTATACATGCAAGTGGAAACGGGACGACGGATCTTGAT 60  
DB 172301 AGTTAAACGCTGGGCGGATGCTTATACATGCAAGTGGGACGACGATGCTTAT 172360  
QY 61 CT-----GCTGGCGAGTGGCGGACGGGTGAGTAAATGATCGGAACGTATCCAGAAAGGGG 116  
DB 172361 AACATGGGTGGCGAGTGGCGGACGGGTGAGTAAATGATCGGAACGTATCGTGGG 172420  
QY 117 GGTAAACGCTGAAAGATGTGCTTAAATCCGCAATATCTTAAAGAGAAAGACGGGATC 176  
DB 172421 GATTAAGTGGGAAACCTTAAGCTTAAATACGATACGACCTTAAGGTAAGAGCGGGGAC 172480  
QY 177 GAAAGCCTTGCGCTTTTGGAGCGGCGGATGCTGATAGCTAGTTGGTGGGGTAAAGG 236  
DB 172481 TTAGGGCCTTGCTGATGATGAGCGGATGCTGATAGCTAGTTGGTGGGGTAAAGG 172540  
QY 237 CTACCAAGGCGATCATGATGTTGTTCTGAGAGAGACGACGACCACTGGGACTGAGA 296

DB 172541 TCACCAAGGCGATCATGATGTTGTTCTGAGAGAGATCATGACGACCACTGGGACTGAGA 172600  
QY 297 CACGGCCGAGACTCTTAAGGAGGCGAGGAGTGGGAAATTTTGGACATGGCGGACGCT 356  
DB 172601 CACGGCTCAGACTCTTAAGGAGGCGAGGAGTGGGAAATTTTGGACATGGCGGACGCT 172660  
QY 357 GATCCAGCAATGCCGCGTGAATGAAGAGGCTTCGGGTTTAAAGCTCTTCACTGAG 416  
DB 172661 GATCCAGCAATGCCGCGTGAATGAAGAGGCTTCGGGTTTAAAGCTCTTGTGGG 172720  
QY 417 AAGAAAGCTTACGTTAAATGCTGATCATGACGCTATGACAGAAAGAGCCGCG 476  
DB 172721 AAGAAAGCTTACGTTAAATGCTGATCATGACGCTATGACAGAAAGAGCCGCG 172780  
QY 477 TAACTACGTGCGAGAGCGCGGTAATACGTAAGGCTGACGCTTAATCGAATTAATCTGG 536  
DB 172781 TAACTACGTGCGAGAGCGCGGTAATACGTAAGGCTGACGCTTAATCGAATTAATCTGG 172840  
QY 537 GCGTAAAGGCTGCGAGAGCGCGCTTGTAAATGCTGATGAAATCCCGGGCTTAACCTGG 596  
DB 172841 GCGTAAAGGCTGCGAGAGCGCGCTTGTAAATGCTGATGAAATCCCGGGCTTAACCTGG 172900  
QY 597 GAATTGCGTTGAAACTCAAGGCTAGAGTGGGAGAGGAGTGAATTCATGTGTA 656  
DB 172901 GAATTGCGTTGAAACTCAAGGCTAGAGTGGGAGAGGAGTGAATTCATGTGTA 172960  
QY 657 GCAGTGAATGCGTGAATGAGATGAGAAACATCGATGCGAAGGCGCTCTGGGTTAA 716  
DB 172961 GCAGTGAATGCGTGAATGAGATGAGAAACATCGATGCGAAGGCGCTCTGGGTTAA 173020  
QY 717 CACTGACGCTCATGCAAGAAACGCTGGGAGCAACAGATTAATCCCTGGTGTCA 776  
DB 173021 CACTGACGCTCATGCAAGAAACGCTGGGAGCAACAGATTAATCCCTGGTGTCA 173080  
QY 777 GCGCTTAAACGATGCAATGCTAGTTGGGCTTATAGG--CTTGATACGAACTAAC 834  
DB 173081 GCGCTTAAACGATGCAATGCTAGTTGGGCTTATAGG--CTTGATACGAACTAAC 173140  
QY 835 GCGTGAATGACCGCTGGGAGTACGCTGCAAGATTAATACTCAAGAAATGACGG 894  
DB 173141 GCGTGAATGACCGCTGGGAGTACGCTGCAAGATTAATACTCAAGAAATGACGG 173200  
QY 895 GGAACCCGCAAGGCGGTGATTAATGATTAATTCGATGCAACCGGAAATCTTACCT 954  
DB 173201 GGAACCCGCAAGGCGGTGATTAATGATTAATTCGATGCAACCGGAAATCTTACCT 173260  
QY 955 ACCCTGACATGACGAATTTTCTAGAGATGATTAAGT-CTTGGGAAACGCTAACACA 1013  
DB 173261 ACCCTGACATGACGAATTTTCTAGAGATGATTAAGT-CTTGGGAAACGCTAACACA 173320  
QY 1014 GGTGCTGACGTGCTGCTGACGCTGCTGAGATGTTGGGTTAAGTCCGCAACGAG 1073  
DB 173321 GGTGCTGACGTGCTGCTGACGCTGCTGAGATGTTGGGTTAAGTCCGCAACGAG 173380  
QY 1074 GCGAACCTTGTGATTAATGCTATCA--TTTGTTGGGCACTTTAATGACCTGCGGT 1131  
DB 173381 GCGAACCTTGTGATTAATGCTATCA--TTTGTTGGGCACTTTAATGACCTGCGGT 173440  
QY 1132 GACAAACCGGAGAGTGGGAGTACGCTCAAGTCTCATAGGCTCTTAATGGGAGCTT 1191  
DB 173441 GACAAACCGGAGAGTGGGAGTACGCTCAAGTCTCATAGGCTCTTAATGGGAGCTT 173500  
QY 1192 CACAGCTAATACATAGGCGCGCTTACAGAGGCTTCCAAACCGGAGGGAGCTTAATCTCA 1251  
DB 173501 CACAGCTAATACATAGGCGCGCTTACAGAGGCTTCCAAACCGGAGGGTGAAGCTTAATCTCA 173560  
QY 1252 GAAAGCGCTGTAAGTCCGATCGAGTCTGCACTGCACTCGTGAAGTCCGAATCGCT 1311  
DB 173561 GAAAGCGCTGTAAGTCCGATCGAGTCTGCACTGCACTCGTGAAGTCCGAATCGCT 173620  
QY 1312 AGTAAATCGGATCAACA-TGTGCGGCTGAATACGTTCCCGGGCTTGTACACACGCGCC 1370

Db 173621 AGTATCGCAGATCGATCGATTGCTGGGTAATAGCTTCCGGGCTTGTGACACCGCC 173680  
Qy 1371 GTACACCATGGAGTGGGTTTACACAGAGCAGGTACTTAACCGTAAAGAGGGCGCTT 1430  
Db 173681 GTACACCATGGAGTGGGTTTGTGACACAGAGCAGGTACTTAACCGCAGAGGGCGCTT 173740  
Qy 1431 GCCACGGTGAATTCATGACTGGGGTG 1457  
Db 173741 GCCACGGTGTGGCCGATGACTGGGGTG 173767

## RESULT 68

US-10-418-861B-53

Sequence 53, Application US/10418861B

Publication No. US20040010131A1

GENERAL INFORMATION:

APPLICANT: da Silva, Ana Claudia Raseira

APPLICANT: Farah, Shaker Chuck

APPLICANT: Ouaguo, Ronaldo Bento

APPLICANT: Reimach, Fernando de Castro

APPLICANT: Ferro, Jesus Aparecido

APPLICANT: De Oliveira, Julio Cezar Franco

APPLICANT: De Lala, Marcelo Luiz

APPLICANT: Secubal Joao C.

APPLICANT: Furian, Luiz Roberto

TITLE OF INVENTION: Isolated Xanthomonas nucleic acid molecules, proteins encoded the

TITLE OF INVENTION: uses thereof

FILE REFERENCE: PAPER 205.1 US

CURRENT APPLICATION NUMBER: US/10/418,861B

CURRENT FILING DATE: 2003-04-17

PRIOR APPLICATION NUMBER: US 60/374,620

PRIOR FILING DATE: 2002-04-22

NUMBER OF SEQ ID NOS: 85

SEQ ID NO 53

LENGTH: 1547

TYPE: DNA

ORGANISM: Xanthomonas

FEATURE:

US-10-418-861B-53

Query Match 70.8%; Score 1031.8; DB 6; Length 1547;

Best Local Similarity 83.7%; Pred. No. 7.5e-265;

Matches 1230; Conservative 0; Mismatches 227; Indels 12; Gaps 5;

Qy 1 ATTTGAACGCTGGCGGCACTTCTTACATGCAAGTCGAAACGGACACGAGATGCTTGAT 60  
Db 28 AGTGAACGCTGGCGGCACTTCTTACATGCAAGTCGAAACGGACACGAGATGCTTGAT 87  
Qy 61 CT-----GGTGGCAGTGGCGGCGGAGTGAAGTAATGATCGGAACGTATCCAGAAAGG 114  
Db 88 CTCTTATGGGTGGCGAGTGGCGGAGTGAAGTAATGATCGGAACGTATCTTCTTCTGG 147  
Qy 115 GGGGTAAAGCATGCAAGATGTCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGA 174  
Db 148 GGGATTAAGTGAAGAACTTACGTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGA 207  
Qy 175 TCGAAAGACCTTGGCTTTTGGAGCGCGCGATGTCTGATTAAGTCTGTTGGGTAAAG 234  
Db 208 CTTTGGGCTTTCGCGCGGTGTAATGAGCGGATGTGGAATTAAGTCTGTTGGGTAAAG 267  
Qy 225 GCTTACCAAGGCGAGATCGATGATGTTGGTCTGAGAGAGCAGACCGACCACTGGAGCTGA 294  
Db 268 GCCCAAGGCGAGATCGATGATGTTGGTCTGAGAGAGTGAATTAAGTGAATTAAGTGA 327  
Qy 295 GACACGCGCAGACTCTTACCGGAGCAGCAGTGGGAAATTTTGAACAATGGGCGCAAGC 354  
Db 328 GACACGCGCAGACTCTTACCGGAGCAGCAGTGGGAAATTTTGAACAATGGGCGCAAGC 387  
Qy 355 CTGATTCAGCAATGCGCGGTGATGTAAGAAGCCTTGGGTGTAAGTCTTTAGTCG 414  
Db 388 CTGATTCAGCAATGCGCGGTGATGTAAGAAGCCTTGGGTGTAAGTCTTTAGTCG 447  
Qy 415 AGAAGAAAAGTTAGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 474

Db 448 GAAAGAAAAGTGGGTAAATACCGATTTGTTGACGGTACCCAAAGAAATTAAGCAGC 507  
Qy 475 GCTAATCTAGTGGCCAGCAGCCCGGTAAATAGTGGGTGCAAGGTTAAATGCAATTA 534  
Db 508 GCTAATCTAGTGGCCAGCAGCCCGGTAAATAGTGGGTGCAAGGTTAAATGCAATTA 567  
Qy 535 GGGGTAAAGGAGTGGCGAGCGCTTGTAAATGATGATGATGATGATGATGATGATGAT 594  
Db 568 GGGGTAAAGGAGTGGCGAGCGCTTGTAAATGATGATGATGATGATGATGATGATGAT 627  
Qy 595 GGGATTTGCGTTTGAATCTAAGCTTGAAGTGTGCAAGAGGAGTGAATTCATGTG 654  
Db 628 GGAATTTGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687  
Qy 655 TAGCAGTAAATGCGTGAAGATTAAGAAAGCAGTATGCGAAGCGACCTCTGGGTT 714  
Db 688 TAGCAGTAAATGCGTGAAGATTAAGAAAGCAGTATGCGAAGCGACCTCTGGAGC 747  
Qy 715 AACACTGACGCTCATGACAGAAAGCGTGGGAGCAAGAGTAAATGATGATGATGATGAT 774  
Db 748 AACACTGACGCTCATGACAGAAAGCGTGGGAGCAAGAGTAAATGATGATGATGATGAT 807  
Qy 775 CACGCGCTTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832  
Db 808 CACGCGCTTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867  
Qy 833 ACCGCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 892  
Db 868 ACCGCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927  
Qy 893 GGGAGCCGCAAGCGGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 952  
Db 928 GGGAGCCGCAAGCGGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 987  
Qy 953 CTACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011  
Db 988 CTGATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1047  
Qy 1012 CAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071  
Db 1048 CAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107  
Qy 1072 AGCGCAACCTTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129  
Db 1108 AGCGCAACCTTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167  
Qy 1130 GTGCAAAACCGAGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGAT 1189  
Db 1168 GTGCAAAACCGAGAGAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGAT 1227  
Qy 1190 TTCAAGTAAATTAAGTGGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 1249  
Db 1228 TTCAAGTAAATTAAGTGGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 1287  
Qy 1250 CAGAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1309  
Db 1288 CAGAAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347  
Qy 1310 CTAGTAATCGGAGTACAGCA-TGTGCGCGTGAATAGTCTTCCGGGTCTTGTACACACCGC 1368  
Db 1348 CTAGTAATCGGAGTACAGCA-TGTGCGCGTGAATAGTCTTCCGGGTCTTGTACACACCGC 1407  
Qy 1369 CCGTCAACCATGAGAGTGGTCTTCAACCAAGAGAGTGAATGATGATGATGATGATGATGAT 1428  
Db 1408 CCGTCAACCATGAGAGTGGTCTTCAACCAAGAGAGTGAATGATGATGATGATGATGATGAT 1467  
Qy 1429 TTGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457  
Db 1468 TTGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496

## RESULT 69

```

US-10-464-724-3
; Sequence 3, Application US/10464724
; Publication No. US2003020398A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; TITLE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BC1033 US NA
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Unknown Organism
; OTHER INFORMATION: Description of Unknown Organism: Denitrifying bacteria
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RAI
US-10-464-724-3

```

```

Query Match 70.4%; Score 1026.4; DB 6; Length 1450;
Best Local Similarity 84.4%; Pred. No. 2e-263;
Matches 1217; Conservative 0; Mismatches 211; Indels 14; Gaps 5;

QY 1 ATTGAACGCTGGCGGCGATCTTACATGCAAGTGGAAAGCGGAGCGGATGCTTGCAT 60
DB 3 ATTGAACGCTGGCGGCGATCTTACATGCAAGTGGAAAGCGGAGCGGATGCTTGCAT 58
QY 61 CTGGTGGCGAGTGGCGGAGCGGAGTGAATGCAATCGGAAAGTATCCAGAAAGAGGGGGG 120
DB 59 CTGGCGGCGAGCGGCGGAGTGAATGCAATCGGAAAGTATCCAGAAAGAGGGGGG 118
QY 121 AGCGATCGAAAGATGCTAATACGCGATATCTTAAGAGGAAAGAGGGGATGAAA 180
DB 119 GCGCGGCGAAAGCGGATTAATACGCGATATCTTAAGAGGAAAGAGGGGATGAAA 178
QY 181 GACCTTGGCGCTTTGGAGCGGCGGATGCTGATTAAGTGGTGGGTTAAAGGCTTAC 240
DB 179 GCGCTCAACGCTTTGGAGCGGCGGATGCTGATTAAGTGGTGGGTTAAAGGCTTAC 238
QY 241 CAAGGCGAGATGATGTTGCTGAGAGAGCAAGCAAGCAAGTGGGAGTGAAGACG 300
DB 239 CAAGCTGCGATCTGATGCTGCTGAGAGATGATCAAGCAAGTGGGAGTGAAGACG 298
QY 301 GCGCAAGTCTTACGCGAGGAGCAAGTGGGAAATTTTGGACATGGGCGGCAAGCTGATC 360
DB 299 GCGCAAGTCTTACGCGAGGAGCAAGTGGGAAATTTTGGACATGGGCGGCAAGCTGATC 358
QY 361 CAGCAATGCGCGGCTGAGTGAAGAGGCTTGGGTTGTAAGCTCTTTCACTGAGAGA 420
DB 359 CAGCAATGCGCGGCTGAGTGAAGAGGCTTGGGTTGTAAGCTCTTTCACTGAGAGA 418
QY 421 AAGGTTACGCTAATATATCTGATCATGACGCTATGACAGAGAGCAAGGCTTAC 480
DB 419 AAGGCTCTCTCTAATATCAAGGCGCATATGACGCTATGAGAAATGAGCAAGGCTTAC 478
QY 481 TAGCTGCGACAGCGCGGCTAATGCTAAGGCTGCAAGCTTAACTGAGGCGT 540
DB 479 TAGCTGCGACAGCGCGGCTAATGCTAAGGCTGCAAGCTTAACTGAGGCGT 538
QY 541 AAGGCTGCGACAGCGGCTTGTATGATGATGTAATCCCGGCTTAACTGAGGAGT 600
DB 539 AAGGCTGCGACAGCGGCTTGTATGATGATGTAATCCCGGCTTAACTGAGGAGT 598
QY 601 TGGCTTTGAAACTGCAAGGCTGAGTGTGCAAGGAGGAGTGAATTCATGTGTACAG 660
DB 599 TGGCTTTGATGCTCAAGGCTGAGTGTGCAAGGAGGAGTGAATTCATGTGTACAG 658

```

```

QY 661 TGAATGCTGATGATATGGAAGAACATTCATGCGAAGGCGAGCTCTGGGTTACACT 720
DB 659 TGAATGCTGATGATATGCGAGGAACCCGATGTAAGGCAATTCCTGGGATGACACT 718
QY 721 GAGCGTCAATGCAAGAAAGCGTGGGAGAGCAACAGGATTAATGCTTGGTGTCCAGCC 780
DB 719 GAGCGTCAATGCAAGAAAGCGTGGGAGAGCAACAGGATTAATGCTTGGTGTCCAGCC 778
QY 781 CT-AAACGATGCTCACTAGTTGTTGGGCTTATTAAGCTTGGTAAAGAACTTACGCGT 839
DB 779 CTAAAGCATGCTCACTAGTTGTTGGGCTTATTAAGCTTGGTAAAGAACTTACGCGT 838
QY 840 AAGTTGACCGCTGGGAGGATACGCTGCGAAGATTTAACTCAAGGATTTGACGGGAGC 899
DB 839 AAGTTGACCGCTGGGAGGATACGCTGCGAAGATTTAACTCAAGGATTTGACGGGAGC 898
QY 900 CGCACAGCGGATGATATGATGATTAATTCGATGCAAGCGAAGAAACCTTACCTACCT 959
DB 899 CGCACAGCGGATGATATGATGATTAATTCGATGCAAGCGAAGAAACCTTACCTACCT 958
QY 960 TGAATGCTGATGATATGGAAGAACATTCATGCGAAGGCGAGCTCTGGGTTACACT 1016
DB 959 TGAATGCTGATGATATGCGAGGAACCGAGAGTCTCGAAGAGAGCGGTAACAGGT 1018
QY 1017 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076
DB 1019 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
QY 1077 AACCTTGCATTAATGCTGATCATTTGGTGGGCACTTTAATGAGACTGCGGTGACA 1136
DB 1079 AACCTTGCATTAATGCTGATCATTTGGTGGGCACTTTAATGAGACTGCGGTGACA 1133
QY 1137 ACCGAGAGAGTGGGAGTGAAGTGAAGTCTGATGCGCTTATGAGGCTTACAC 1196
DB 1134 ACCGAGAGAGTGGGAGTGAAGTGAAGTCTGATGCGCTTATGAGGCTTACAC 1193
QY 1197 GTAATCAATGCGCGGTACAGAGGTTGGCAACCCCGAGGGGAGCTAATCTCAGAA 1256
DB 1194 GTATATCAATGCGCGGTACAGAGGTTGGCAACCCCGAGGGGAGCTAATCTCAGAA 1253
QY 1257 CGCGTGTAGTGCAGATGCGGATGCGAGTCTGCAACTGCACTCGGTGAAGTCTAGTAA 1316
DB 1254 CAGTGTGATGCTGCGATGCGGATGCGGATCTGCAACTGCACTCGGTGAAGTCTAGTAA 1313
QY 1317 TCGCGATCAGATGCTGCGGTAATGCTT-CGCGGCTTGTATCAACCGCCGCTAC 1375
DB 1314 TCGTGTATCAGATGCTGCGGTAATGCTT-CGCGGCTTGTATCAACCGCCGCTAC 1373
QY 1376 ACCATGAGAGTGGTTTCAAGAGAGTGTCTAACCTGTAAGAGGCGCTTCCAC 1435
DB 1374 ACACATGAGAGGTTTCTCCAGAAAGTGTAGCTTAACCGAAGAGGCGGATTAACAC 1433
QY 1436 GG 1437
DB 1434 GG 1435

```

```

RESULT 70
US-10-464-356-3
; Sequence 3, Application US/10464356
; Publication No. US20030207320A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael
; TITLE OF INVENTION: Nucleic Acid Fragments for the Identification of Bacteria in
; TITLE OF INVENTION: Industrial Wastewater Bioreactors
; FILE REFERENCE: BC1033 US NA
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US/09/735,567
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/171,140
; PRIOR FILING DATE: 16 DECEMBER 1999

```

```

; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; DESCRIPTION: Description of Unknown Organism: Denitrifying bacteria
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of 16S rDNA corresponding to strain RAI
US-10-464-356-3

```

Query Match 70.4%; Score 1026.4; DB 6; Length 1450;  
 Best Local Similarity 84.4%; Pred. No. 28-263;  
 Matches 1217; Conservative 0; Mismatches 211; Indels 14; Gaps 5;

```

QY 1 ATGTAACGCTGGCGGCGATGCTTTACATGCAATGCAAGCGGACGACGAGATGCTTGCAAT 60
DB 3 ATGTAACGCTGGCGGCGATGCTTTACATGCAATGCAAGCGGACGACGAGATGCTTGCAAT 58
QY 61 CTGGTGGCGAGTGGCGGACGGGTGAGTAATGCAATCGGAACTATCCAGAAAGGCGGGTA 120
DB 59 CTGGCGGCGAGCGCGGACGGGTGAGTAATGCAATCGGAACTATCCAGAAAGGCGGGTA 118
QY 121 AGCATTCGAAAGATGTCCTAATACCGCATATCTCTAAGAGAGAAAGCAGGGGATCGAA 180
DB 119 GCCCGGCGAAAGCCGATTTAATACCGCATGATCTGAGGATGAAAGTGGGGACCGCA 178
QY 181 GACCTTGGCGCTTTTGAAGCGCGCATGCTGATTTAGCTAGTGGTGGGTAAGGCTTAC 240
DB 179 GACCTTACGCGCTTTTGAAGCGCGCATGCTGATTTAGCTAGTGGTGGGTAAGGCTTAC 238
QY 241 CAAGCGCATGATGATGATGCTGAGAGACGACCACTGAGTGGGATGAGACAG 300
DB 239 CAAGCTCGCATGATGATGCTGAGAGATGATGACCCACTGAGTGGGATGAGACAG 298
QY 301 GCCCGATCTCTTACCGGAGGCGACAGTGGGGAATTTTGGACAAATGGCGGCAAGCTTATC 360
DB 299 GCCCGATCTCTTACCGGAGGCGACAGTGGGGAATTTTGGACAAATGGCGGCAAGCTTATC 358
QY 361 CAGCAATGCGCGGCTGATGTAAGAAAGGCTTCCGGTTGTAAGCTTTTCACTGAGAAAGA 420
DB 359 CAGCAATGCGCGGCTGATGTAAGAAAGGCTTCCGGTTGTAAGCTTTTCACTGAGAAAGA 418
QY 421 AAAGGTTACGCTAATTAATGCTGATCATGACGCTATGCAAGAGAAAGACCGGCTAAC 480
DB 419 AAAGGCTCTCTCTAATTAATGCTGATCATGACGCTATGCAAGAGAAATTAAGACCGGCTAAC 478
QY 481 TACCTGCGACGACGCGCGGTAATACGTAGGGTGCAAGGCTTAATGGAATTAATGCGGCT 540
DB 479 TACCTGCGACGACGCGCGGTAATACGTAGGGTGCAAGGCTTAATGGAATTAATGCGGCT 538
QY 541 AAAGGCTGCGACGCGGCTTTTGAATGCAATGTAATATCCCGGGCTTAACTTGGGAAT 600
DB 539 AAAGGCTGCGACGCGGCTTTTGAATGCAATGTAATATCCCGGGCTTAACTTGGGAAT 598
QY 601 TGGCTTTGAATACTAAGGCTAGAGTGTGGCAGAGGAGAGTGAATTCATGTTAGCAG 660
DB 599 TGGCTTTTGTGATGCAAGGCTGAGTGGCGCAGAGGGGATGGAATTCGCGCTGAGCAG 658
QY 661 TGAATTTGCTAGAGTATGGAAGAACTGATGGGGAAGGCAATCCCGTGGGCTTAACT 720
DB 659 TGAATTTGCTAGAGTATGGAAGAACTGATGGGGAAGGCAATCCCGTGGGCTTAACT 718
QY 721 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGCTAGTCCAGCGC 780
DB 719 GACGCTCATGACGAAAGCGTGGGAGCAAAACAGATTAGATACCTGCTAGTCCAGCGC 778
QY 781 CT-AAACGATGCAACTAGTTGTTGGGCTTTATAGCTTGGTAAAGAACTAACGCGTG 839
DB 779 CTAAACGATGCAACTAGTTGTTGGGCTTTATAGCTTGGTAAAGAACTAACGCGTG 838
QY 840 AAGTTGACCGCTGGGAGTACGCTGCAAGATTAAACTCAAGAAATGACGGGAGC 899

```

```

DB 839 AAGTTGACCGCTGGGAGTACGCTGCAAGATTAAACTCAAGAAATGACGGGAGC 898
QY 900 CGCAACAGCGGTGATTAATGATTAATTCGATGCAACGGAAGAACTTACTACCT 959
DB 899 CGCAACAGCGGTGATTAATGATTAATTCGATGCAACGGAAGAACTTACTACCT 958
QY 960 TGACATGAGCGAATTTTCTAGATAGTATAGTCT---TCGGGAACGCTAACACAGGT 1016
DB 959 TGACATGAGCGAATTTTCTAGATAGTATAGTCT---TCGGGAACGCTAACACAGGT 1018
QY 1017 GCTGCATGCTGTGCTGACGCTGCTGCTGAGATGTTGGTAAAGTCCCGCAAGAGCGC 1076
DB 1019 GCTGCATGCTGTGCTGACGCTGCTGCTGAGATGTTGGTAAAGTCCCGCAAGAGCGC 1078
QY 1077 AACCTTGTCAATTAATGCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1136
DB 1079 AACCTTGTCAATTAATGCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
QY 1137 ACCGAGGAAGTGGGATGATGCTCAAGTCTCTCATGCTGCTTATGAGTGGGCTTCAAC 1196
DB 1134 ACCGAGGAAGTGGGATGATGCTCAAGTCTCTCATGCTGCTTATGAGTGGGCTTCAAC 1193
QY 1197 GTAATCAATGCGCGCTGACAGAGGTTGCCAACCCGCGAGGGGAGCTTAATCTAGAAAG 1256
DB 1194 GTCATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1253
QY 1257 GCGCTGCTGATGCTGCGATGCGAGTCTGCAACTGCACTCGTGAATCGGAATCGCTAGTAA 1316
DB 1254 CCAATGCTGATGCTGCGATGCGAGTCTGCAACTGCACTCGTGAATCGGAATCGCTAGTAA 1313
QY 1317 TCGGAGATCAAGATGCTGCGGTTGAATAGTT---CCCGGCTTGTGACACACCGCCGCTAC 1375
DB 1314 TCGGAGATCAAGATGCTGCGGTTGAATAGTT---CCCGGCTTGTGACACACCGCCGCTAC 1373
QY 1376 ACCATGGAAGTGGGTTTCAACGAGAGAGTGTAACTTAAGTAAAGAGGCGCTTGGCAC 1435
DB 1374 ACACATGAGAGGCTTCTGCGAAGATGTTAGCTTAACGCAAGAGGAGGATTAACAC 1433
QY 1436 GG 1437
DB 1434 GG 1435

```

Search completed: May 5, 2006, 14:07:23  
 Job time : 1293 secs